

## **Primer and database choice affect fungal functional but not biological diversity findings in a national soil survey**

George, Paul; Creer, Simon; Griffiths, Robert I.; Emmett, Bridgett A.; Robinson, David A.; Jones, Davey L.

**Frontiers in Environmental Science**

DOI:

[10.3389/fenvs.2019.00173](https://doi.org/10.3389/fenvs.2019.00173)

Published: 01/11/2019

Peer reviewed version

[Cyswllt i'r cyhoeddiad / Link to publication](#)

*Dyfyniad o'r fersiwn a gyhoeddwyd / Citation for published version (APA):*

George, P., Creer, S., Griffiths, R. I., Emmett, B. A., Robinson, D. A., & Jones, D. L. (2019). Primer and database choice affect fungal functional but not biological diversity findings in a national soil survey. *Frontiers in Environmental Science*, 7, [173].  
<https://doi.org/10.3389/fenvs.2019.00173>

### **Hawliau Cyffredinol / General rights**

Copyright and moral rights for the publications made accessible in the public portal are retained by the authors and/or other copyright owners and it is a condition of accessing publications that users recognise and abide by the legal requirements associated with these rights.

- Users may download and print one copy of any publication from the public portal for the purpose of private study or research.
- You may not further distribute the material or use it for any profit-making activity or commercial gain
- You may freely distribute the URL identifying the publication in the public portal ?

### **Take down policy**

If you believe that this document breaches copyright please contact us providing details, and we will remove access to the work immediately and investigate your claim.

**Primer and database choice affect fungal functional but not biological diversity  
findings in a national soil survey**

Paul B.L. George<sup>\*,1,2</sup>, Simon Creer<sup>1</sup>, Robert I. Griffiths<sup>2</sup>, Bridget A. Emmett<sup>2</sup>, David A.  
Robinson<sup>2</sup>, Davey L. Jones<sup>1,3</sup>

<sup>1</sup> *School of Natural Sciences, Bangor University, Deiniol Road, Bangor, Gwynedd, LL57  
2UW, United Kingdom*

<sup>2</sup> *Centre for Ecology & Hydrology, Environment Centre Wales, Deiniol Road, Bangor,  
Gwynedd, LL57 2UW, United Kingdom*

<sup>3</sup> *School of Agriculture and Environment, University of Western Australia, Crawley,  
Western Australia, WA6009, Australia*

\*Corresponding author. School of Natural Sciences, Bangor University, Deiniol Road,  
Bangor, Gwynedd, LL57 2UW, United Kingdom. Email: [afp67e@bangor.ac.uk](mailto:afp67e@bangor.ac.uk);  
[pblgeorge@gmail.com](mailto:pblgeorge@gmail.com)

**British English formatting**

**KEYWORDS:** UNITE; SILVA; identification bias; high-throughput sequencing;  
Arbuscular mycorrhizal fungi; Archaeorhizomycetes

## ABSTRACT

The internal transcribed spacer (ITS) region is the accepted DNA barcode of fungi. Its use has led to a step-change in the assessment and characterisation of fungal communities from environmental samples by precluding the need to isolate, culture, and identify individuals. However, certain functionally important groups, such as the arbuscular mycorrhizas (Glomeromycetes), are better characterised by alternative markers such as the 18S rRNA region. Previous use of an ITS primer set in a nationwide metabarcoding soil biodiversity survey revealed that fungal richness declined along a gradient of productivity and management intensity. Here, we wanted to discern whether this trend was also present in data generated from universal 18S primers. Furthermore, we wanted to extend this comparison to include measures of functional diversity and establish trends with soil types and soil organic matter (SOM) content. Over the 413 individual sites examined (arable, grassland, woodland, moorland, heathland), we found congruent trends of total fungal richness and  $\beta$ -diversity across land uses, SOM class and soil type with both ITS and 18S primer sets. A total of 24 fungal classes were shared between datasets, in addition to 15 unique to ITS1 and 12 unique to 18S. However, using FUNGUILD, divergent trends of functional group richness became apparent, especially for symbiotrophic fungi, likely driven by an increased detection rate of Glomeromycetes in the 18S dataset. The disparate trends were also apparent when richness and  $\beta$ -diversity were compared to soil properties. Additionally, we found SOM class to be a more meaningful variable than soil type biodiversity for predicting biodiversity analyses because organic matter was calculated for each sample whereas soil type was assigned from a national soil map. We advocate that a combination of fungal primers should be

used in large-scale soil biodiversity surveys to capture important groups that can be underrepresented by universal barcodes. Utilising such an approach can prevent the oversight of ubiquitous but poorly described species as well as critically important functional groups.

## INTRODUCTION

Soil fungi are the dominant eukaryotic component of soil communities and are known to perform crucial ecosystem functions (Peay et al., 2008). Characterising the diversity of fungi within the landscape and their response to anthropogenic perturbation therefore represents an important topic within ecology. High-throughput sequencing has allowed the rapid estimation and identification of fungi by overcoming historical limitations of culture isolation and classifying fruiting bodies (Tedersoo et al., 2015). Using these DNA-based approaches it has been estimated that global fungal diversity in soil ranges from 3.5 – 5 million species. Yet at the beginning of the present decade, only around one-tenth of fungal diversity was thought to have been described (Rosling et al., 2011). In terms of ecosystem function, the majority of fungi are important in organic matter turnover and nutrient recycling as they facilitate the conversion of complex organic polymers into forms more readily accessible to other organisms (Peay et al., 2008; Nguyen et al., 2016). Consequently, they play a crucial role in regulating both below- and above-ground productivity (Peay et al., 2008). Many soil fungi also form important interactions with plants. Some form mutualistic relationships, best exemplified by the wide range of mycorrhizas (Wang and Qui, 2006; Smith and Read, 2008; Nguyen et al., 2016), whereas others are pathogens, responsible for numerous plant and animal

diseases within agriculture and forestry (Fisher et al., 2012; Nguyen et al., 2016). Depending on environmental conditions or life stage, fungi are capable of taking on some or all of these roles (i.e. saprotroph, symbiotroph, pathotroph) (Fisher et al., 2012). Despite the recognition that fungi are extremely important in soil ecosystems, characterising fungal communities has remained a challenge, exemplified by the numerous studies on soil bacteria in comparison to fungi.

Fungal barcode sequences are found within the ubiquitous, multicopy ribosomal RNA gene. Within this, the internal transcribed spacer (ITS) region has been accepted as a universal barcode for fungi (Schoch et al., 2012). Recent development of ITS-based databases such as UNITE (Kõjlg et al., 2013) and Warcup (Deshpande et al., 2016) have overcome limitations in collecting and assigning taxonomic identities to unknown sequences, though database selection may introduce bias into results (Tedersoo et al., 2015; Xue et al., 2019). Yet ITS barcodes exhibit some limitations when dealing with unknown or environmental samples. Generally, the ITS region cannot be aligned above the family-level (Cavender-Bares et al., 2009), making phylogenies based on ITS sequence data unreliable. Importantly, the ITS region has proven unreliable at distinguishing certain fungal groups at the species-level, such as Glomeromycetes (Stockinger et al., 2010). Such inconsistencies mean that ITS primers may not accurately detect target organisms. For instance, Berruti et al. (2017), found that ITS primers underestimated Glomeromycetes in bulk soil. Such uncertainty may confound experimental results and lead to erroneous conclusions.

Despite the widespread use of ITS barcodes, other markers may better capture the diversity of some fungal taxa. Primers targeting the small and large subunits as well as

the ITS regions of the rRNA gene have all been applied to fungi (Tedersoo et al., 2015; Xue et al., 2019). For example, early diverging lineages such as Chytridiomycota (Schoch et al., 2012; Tedersoo et al., 2015) and Glomeromycetes (Tedersoo, et al., 2015) are poorly represented in ITS sequencing. Additionally, advancements in classification have highlighted the shortcomings of environmental DNA barcoding. For example, the Archaeorhizomycetes are a poorly understood but ubiquitous class of soil fungi and their previously unidentifiable sequences have been major components of past soil biodiversity assessments (Anderson et al., 2003; Rosling, et al., 2011). Overlooking these lineages may potentially lead to erroneous assumptions of biological and functional diversity in soils.

Underrepresentation of Glomeromycetes in particular exemplifies this issue. Arbuscular mycorrhizal fungi (AMF) form symbiotic relationships with more than 80% of vascular plant families and have been categorised into the monophyletic Glomeromycetes (Schüßler et al., 2001). Unlike most fungi, the ITS region has consistently demonstrated poor resolution in some closely related AMF species (Stockinger et al., 2010) as it is too hyper-variable (Thiéry et al., 2016). As mentioned previously, the ITS region underestimates Glomeromycetes in bulk soil (Berruti et al., 2017). Instead, the 18S region is more commonly used for barcoding AMF, especially in ecological studies (Öpik et al., 2014). Therefore it is important to recognise biases inherent even in supposedly universal barcodes.

We previously undertook a nation-wide assessment of soil biodiversity across Wales, representing a breadth of heterogeneous land uses, which included agricultural land, grasslands, woodlands, and upland bogs. In this case, fungal richness and  $\beta$ -

diversity were assessed using soil environmental DNA, utilising ITS1 primers (George et al., 2019). Yet, from the earliest stages of experimental design, we were cognisant that the ITS1 universal primer choice may not account for numerous functionally important fungal groups, particularly AMF. Thus, the primary objective of the present study was to assess whether observed fungal biodiversity (richness and  $\beta$ -diversity) across contrasting land uses from the ITS1 dataset would differ when compared to a dataset derived from an alternative choice of primer and database. We therefore sought to assess if primer choice influenced fungal biodiversity across land use, soil type, and soil organic matter (SOM) class. Our next aim was to critically evaluate the influence of climatic and edaphic factors (e.g. soil pH, total carbon (C), nitrogen (N), phosphorus (P)) on fungal diversity arising from the use of the two different primer sets. Our final aim was to look for differences in coverage of taxonomic and functional diversity between the two primer sets across the broad range of land uses and soil types evaluated.

## **MATERIALS AND METHODS**

### *Study Design*

Data were collected as part of the Glastir Monitoring & Evaluation Programme (GMEP). The GMEP initiative was established by Welsh Government to monitor their most recent agri-environment scheme, Glastir, which involved 4,911 landowners over an area of 3,263 km<sup>2</sup> (Fig. 1). Through the GMEP framework, survey teams collected samples in 2013 and 2014 between April and October in each year (Emmett and the GMEP Team, 2017). Sampling protocols were based on those of the UK-wide ecosystem monitoring programme, Countryside Survey (Emmett et al., 2010). The survey design

randomly located 300, 1 km squares across 26 land classes in Wales which survey teams sampled with 5 plots in each square. A subset of samples were then randomly chosen from squares with a maximum of 3 selected in an individual square. A total of 437 samples were collected for biodiversity analyses.

At each sampling location, 2 cores were collected. One was a 15 cm deep by 4 cm diameter core from which measurements of soil physical and chemical properties were taken, including total C (%), N (%), P (mg/kg), organic matter (% loss-on-ignition), pH (measured in 0.01 M CaCl<sub>2</sub>), mean soil water repellency (water drop penetration time in seconds), bulk density (g/cm<sup>3</sup>), volume of rocks (cm<sup>3</sup>), volumetric water content (m<sup>3</sup>/m<sup>3</sup>), as well as percentage sand and clay. For complete details on chemical analyses methodology, see Emmett et al. (2010). Soil texture data were measured by laser granulometry with a LS320 13 analyser (Beckman-Coulter) as described in George et al. (2019). The cut-off points for clay, silt, and sand were: 2.2 µm, 63 µm and 2000 µm respectively. Clay and sand percentages were selected for subsequent analyses and normalised using Aitchison's log<sub>10</sub>-ratio transformation. Further geographic data including grid eastings, northings, and elevation were also collected. Mean temperature (°C) on date of sample collection and annual precipitation (mL) data were extracted from the Climate Hydrology and Ecology research Support System dataset (Robinson et al., 2017). Environmental variables were normalised (by log<sub>10</sub> or square root transformation) where appropriate (see Table 1).

Each sampling site was assigned to a land use category, soil type, and SOM class (based on percentage organic matter). The land use classification used in this study was originally developed for the UK Countryside Survey in 1990 (Bunce et al., 1999).



Briefly, vegetation was recorded by surveyors and used to classify each site into one of the 8 Aggregate Vegetation Classes (AVCs) as described in Bunce et al. (1999; for further details please see Supplementary Material). The AVCs have been shown to follow a gradient of soil nutrient content from which productivity and management intensity can also be inferred (see Supplementary Material and Bunce et al., 1999). There were 7 AVCs identified in the present study. The AVCs in descending order of productivity are: Crops/weeds (including arable land), Fertile grassland, Infertile grassland, Lowland woodland, Upland woodland, Moorland grass-mosaic, Heath/bog (Supplementary Table 1). Soil type based on the predominant major soil group classification was extracted from the National Soil Map (Supplementary Material; Avery, 1980). Additionally, we classified soils on a per sample basis by organic matter content. Each sample was grouped into one of four organic matter classes based on percent loss-on-ignition (LOI) following the protocols of the 2007 Countryside Survey (Emmett et al., 2010): mineral (0-8% LOI), humus-mineral (8-30% LOI), organo-mineral (30-60% LOI), and organic (60-100% LOI). Mean values for each environmental variable were recorded for each land use, soil organic matter class, and soil type.

#### *DNA Extraction*

Soils used in DNA extraction were collected from 15 cm deep by 8 cm diameter cores. Soil samples were transported in refrigerated boxes; samples were received at Environment Centre Wales, Bangor within an average of 48 h post-extraction and frozen at -80 °C upon arrival. Soils were then thawed and homogenised as they passed through a sterilised 2 mm stainless steel sieve after which they were returned to a -80 °C freezer until DNA extraction. Sieves were sterilised between samples by rinsing with tap water at

high pressure and an application of Vircon<sup>®</sup> laboratory disinfectant followed by UV-  
treating each side for 5 minutes. DNA was extracted by mechanical lysis from 0.25 g of  
soil per sample using a PowerLyzer PowerSoil DNA Isolation Kit (MO-BIO Inc.). Soils  
were pre-treated with 750 µL of a suspension of CaCO<sub>3</sub> (1 M) following Sagova-  
Mareckova et al. (2008) to improve PCR performances, especially for acidic soils.  
Extracted DNA was stored at -20 °C until amplicon library preparation began. The  
extractions and homogenisation steps were performed in triplicate. To check for  
contamination in sieves, 3 negative control DNA extractions were completed as well as 2  
negative control kit extractions using the same technique but without the CaCO<sub>3</sub> pre-  
treatment. Aliquots of the resultant DNA were used to create amplicon libraries for  
sequencing with each primer set.

#### *Primer Selection and PCR Protocols for Library Preparation*

Amplicon libraries were created using primers for the ITS1 (ITS5/5.8S\_fungi)  
area to specifically target fungi (Epp et al., 2012) and the V4 region of the 18S gene  
(TAREuk454FWD1/TAREukREV3) (Behnke et al., 2011) targeting a wide range of, but  
not all, eukaryotic organisms, including fungi. A two-step PCR following protocols  
devised in conjunction with the Liverpool Centre for Genome Research was used as  
described in George et al. (2019). Amplification of amplicon libraries was run in  
triplicate on DNA Engine Tetrad<sup>®</sup> 2 Peltier Thermal Cycler (BIO-RAD Laboratories Inc.)  
and thermocycling parameters for both PCR protocols started with 98 °C for 30 s and  
terminated with 72 °C for 10 min for final extension and held at 4 °C for a final 10 min.  
For the ITS1 locus, there were 15 cycles of 98 °C for 10 s; 58 °C for 30 s; 72 °C for 30 s.

For the 18S locus there were 15 cycles at 98 °C for 10 s; 50 °C for 30 s; 72 °C for 30 s. Twelve µL of each first-round PCR product were mixed with 0.1 µL of exonuclease I, 0.2 of µL thermosensitive alkaline phosphatase, and 0.7 µL of water and cleaned in the thermocycler with a programme of 37 °C for 15 min and 74 °C for 15 min and held at 4 °C. Addition of Illumina Nextera XT 384-way indexing primers to the cleaned first round PCR products were amplified following a single protocol which started with initial denaturation at 98 °C for 3 min; 15 cycles of 95 °C for 30 s; 55°C for 30 s; 72 °C for 30 s; final extension at 72 °C for 5 min and held at 4 °C. Twenty-five µL of second-round PCR products were purified with an equal amount of AMPure XP beads (Beckman Coulter). Library preparation for the 2013 samples was conducted at Bangor University. Illumina sequencing for both years and library preparation for 2014 samples were conducted at the Liverpool Centre for Genome Research.

## *Bioinformatics*

Bioinformatics analyses were performed on the Supercomputing Wales cluster as previously described in George et al. (2019). A total of 104,276,828, and 98,999,009 raw reads were recovered from the ITS1 and 18S sequences, respectively. Illumina adapters were trimmed from sequences using Cutadapt (Martin, 2011) with 10% level mismatch for removal. Sequences were then de-multiplexed, filtered, quality-checked, and clustered using a combination of USEARCH v. 7.0 (Edgar, 2010) and VSEARCH v. 2.3.2 (Rognes et al., 2016). Open-reference clustering (97% sequence similarity) of operational taxonomic units (OTUs) was performed using VSEARCH; all other steps were conducted with USEARCH. Sequences with a maximum error greater than 1 and shorter than 200

bp were removed following the merging of forward and reverse reads for ITS1 sequences. A cut-off of 250 bp was used for 18S sequences, according to higher quality scores. There were 7,242,508 (ITS1) and 9,163,754 (18S) cleaned reads following these steps. Sequences were sorted and those that only appeared once in each dataset were removed.

Remaining sequences were matched first against the UNITE 7.2 (Kõljalg et al., 2013) and SILVA 128 (Quast et al., 2013) databases for the ITS1 and 18S sequences, respectively. Ten per cent of sequences that failed to match were clustered *de novo* and used as a new reference database for failed sequences. Sequences that failed to match with the *de novo* database were subsequently also clustered *de novo*. All clusters were collated and chimeras were removed using the uchime\_ref command in VSEARCH. Chimera-free clusters and taxonomy assignment summarised in an OTU table with QIIME v. 1.9.1 (Caporaso et al., 2010) using RDP (Wang et al., 2007) methodology with the UNITE database for ITS1 data. Taxonomy was assigned to the 18S OTU table using BLAST (Altschul et al., 1990) against the SILVA database and OTUs appearing only once or in only 1 sample were removed from each OTU table. Based on DNA quality and read counts, 413 samples were used for analyses of the ITS1 data and 422 for 18S data (from the total of 438).

A Newick tree was constructed for the 18S tables using 80% identity thresholds and was paired with the 18S OTU table as part of analyses using the R package phyloseq (McMurdie and Holmes, 2013). Non-fungi OTUs were removed from both OTU tables. Read counts from each group were rarefied 100 times using phyloseq (as justified by Weiss et al. (2017)) and the resulting mean richness was calculated for each sample. The

ITS1 table was rarefied at a depth of 4,000 reads whereas the 18S table was rarefied to 10,000 reads. A subset of the 18S data was rarefied to 400 reads across 398 samples to analyse Glomeromycetes OTUs separately. Samples with observed lower read counts were removed before rarefaction. To assess functional diversity, both OTU tables were processed using FUNGUILD (Nguyen et al., 2016) and the resulting matched OTU tables were used to investigate functional roles based on trophic mode. Sequences have been uploaded to The European Nucleotide Archive and can be accessed with the following primary accession codes after the end of the data embargo: PRJEB28028 (ITS1), and PRJEB28067 (18S).

#### *Statistical Analysis*

All statistical analyses were run using R v. 3.3.3 (R Core Team, 2017) following rarefaction. For each data set, NMDS ordinations using Bray-Curtis dissimilarity were created with the vegan package (Oksanen et al., 2016) to assess  $\beta$ -diversity.

Environmental data was fitted linearly onto each ordination of AVCs using the envfit function. NMDS scores were plotted against these values for each variable to determine the direction of associations. Differences in  $\beta$ -diversity amongst AVCs were calculated with PERMANOVA and homogeneity of dispersion was also assessed.

Linear mixed models were constructed using package nlme (Pinheiro et al., 2016) to show the differences in  $\alpha$ -diversity amongst AVCs, soil types, and LOI classification, for both ITS1 and 18S fungal data sets. Sample year as fixed factors; sample square identity was the random factor. This methodology was also used for the subsets of data that matched to the FUNGUILD database. For each model, significant differences were

assessed by ANOVA and pairwise differences were identified using Tukey's *post-hoc* tests from the multcomp package (Hothorn et al., 2008).

Partial least squares regressions from the pls package (Mevik et al., 2016) were used with the variable importance in projection (VIP) approach (Chong and Jun, 2005) to sort the original explanatory variables by order of importance to identify the most important environmental variables for richness. Such analysis is ideal for data where there are many more explanatory variables than sample numbers or where extreme multicollinearity is present (Lallias et al., 2015; George et al., 2019). Variables with VIP values  $> 1$  were considered most important. Relationships between important variables and richness values for each group of organisms were investigated by linear regression. Richness was normalised before regression when necessary.

## RESULTS

### *Soil Properties*

Soil properties displayed a range of changes across land uses (Table 1). Notably, total C ( $F_{6, 427} = 89.13$ ,  $p < 0.001$ ), total N ( $F_{6, 427} = 61.03$ ,  $p < 0.001$ ), C:N ratio ( $F_{6, 427} = 94.41$ ,  $p < 0.001$ ), organic matter content ( $F_{6, 428} = 107.02$ ,  $p < 0.001$ ), elevation ( $F_{6, 429} = 78.42$ ,  $p < 0.001$ ), and mean annual precipitation ( $F_{6, 429} = 72.6$ ,  $p < 0.001$ ), and moisture ( $F_{6, 427} = 33.74$ ,  $p < 0.001$ ) increased with declining land use productivity. We also observed a reduction in pH ( $F_{6, 428} = 69.56$ ,  $p < 0.001$ ), bulk density ( $F_{6, 428} = 79.87$ ,  $p < 0.001$ ), and clay content ( $F_{6, 344} = 19.54$ ,  $p < 0.001$ ) across the land use productivity gradient. Trends in other variables such as soil water repellency ( $F_{6, 428} = 22.08$ ,  $p < 0.001$ ), total P ( $F_{6, 424} = 7.1$ ,  $p < 0.001$ ), sand content ( $F_{6, 344} = 5.71$ ,  $p < 0.001$ ), stone

content ( $F_{6, 427} = 10.4$ ,  $p < 0.001$ ), and temperature at time of sampling ( $F_{6, 429} = 4.4$ ,  $p < 0.001$ ), though significant, were less clear across land uses however. These findings were also apparent when samples were grouped from low-to-high organic matter content by organic matter class (Supplementary Table 2). Overall, no clear trends were evident across the different soil types (Supplementary Table 3).

#### *Sequencing Data*

A total of 7,582 and 4,408 fungal OTUs were recovered using the ITS1 and 18S primer sets, respectively. Of these, 5,666 were assigned an identifier at the class-level in the ITS1 dataset while 4,367 were assigned an identifier in the 18S dataset. There were 15 classes that were only found in the ITS1 dataset and 12 unique to the 18S data. Endogonomycetes was the most abundant class found only in the ITS dataset (19 OTUs), whereas Laboulbeniomycetes (17 OTUs) was the most abundant fungal class unique to the 18S data. A total of 24 classes were present in both ITS1 and 18S data (Fig. 2A).

As reported in George et al. (2019), Agaricomycetes were the most abundant class of fungi in the ITS1 dataset overall. There were also a large proportion of Sordariomycetes (Fig. 2B). Archaeorhizomycetes was the most abundant class in the 18S dataset (Fig. 2C). Proportionate abundances of Sordariomycetes and Agaricomycetes followed contrasting trends, with the dominance of the former replaced by the latter in lower productivity AVCs in the ITS1 data, as described previously (Fig. 3A). Although Agaricomycetes and Sordariomycetes comprised smaller fractions of the 18S dataset (Fig. 2C), this trend was still apparent (Fig. 3B). Additionally, the Archaeorhizomycetes from 18S data generally followed the same trend as the Sordariomycetes (Fig. 3B). The

preceding trends observed across land uses are also evident across organic matter classes (Fig. S1) but are not as clear across soil types (Fig. S2).

When a class was present in both datasets, it was usually much more prevalent in one than the other (Supplementary Table 4). For example, there were 1858 Agaricomycetes and 915 Sordariomycetes OTUs in the ITS1, yet these numbers dropped to 646 and 417 OTUs in the 18S dataset. Similarly, Glomeromycetes accounted for 162 of the OTUs in the 18S data, but only 6 OTUs in the ITS1 dataset. Abundances of classes unique to the ITS1 and 18S datasets can be found in Supplementary Table 5 and Supplementary Table 6, respectively.

#### *Fungal Richness and $\beta$ -Diversity from ITS1 and 18S Data*

We found that fungal richness followed the same trends across land use, irrespective of primer set. As previously demonstrated in George et al. (2019), fungal OTU richness from ITS1 metabarcoding significantly declined ( $F_{6, 258} = 39.87$ ,  $p < 0.001$ ; Fig. 4A) from high to low productivity/management intensity. Richness in Fertile grasslands was significantly greater than all other AVCs ( $p < 0.001$ ) except Crops/weeds. In the 18S dataset, richness was also significantly higher ( $F_{6, 267} = 82.73$ ,  $p < 0.001$ ) in more productive/managed land uses and declined along this gradient. However, richness in grasslands was highest in this dataset (Fig. 4B). For complete pairwise differences between land uses see Supplementary Material.

The trend of declining richness with productivity was also apparent when samples were categorised by organic matter content (Fig. 5). In both datasets, richness was significantly greater ( $F_{3, 259} = 48.13$ ,  $p < 0.001$ ;  $F_{3, 269} = 46.71$ ,  $p < 0.001$ ; for ITS1 and



18S, respectively) in mineral and humus-mineral than all other classifications (ITS1, Fig. 5A; 18S, Fig. 5B). There was no consistent pattern of richness when soils were categorised by soil type (Fig. S3). Again pairwise differences between organic matter classes and soil types are described in the Supplementary Material.

Community composition based on non-metric multidimensional scaling of Bray-Curtis distances also showed consistent trends between the datasets. Plots demonstrate tight clustering of Crops/weeds, and grassland AVCs in both ITS1 (Fig. 6A) and 18S (Fig. 6B) compared to the wide dispersal of other AVCs. Such results are supported by PERMANOVAs, which show significant differences ( $F_{6, 406} = 10.74$ ,  $p = 0.001$ ;  $F_{6, 415} = 15.65$ ,  $p = 0.001$ ); however, analyses of dispersion were also significant ( $F_{6, 406} = 41.30$ ,  $p = 0.001$ ;  $F_{6, 415} = 10.69$ ,  $p = 0.001$ ) as a result of the large disparity in replicates between land uses.

When these results are visualised by organic matter classification, the tight clusters are populated by mineral and humus-mineral samples, whereas organo-mineral and organic samples are more common in the widely dispersed areas of the plots (Fig. S4 and Fig. S5). Soil types are more widely dispersed but Brown and Surface-water gley soils are more common in the tightly grouped area (Fig. S6 and Fig. S7). Again, significant results were observed for both PERMANOVA and dispersion of variance across organic matter classes and soil types in both datasets.

#### *Relationships Between Soil Properties and Fungal Biodiversity*

Fungal richness showed similar relationships to soil properties in both datasets. Across samples, PLS and VIP analyses highlighted strong correlations between fungal

richness and soil properties. There were significant, positive relationships of richness with pH and bulk density; and significant, negative correlations between richness and C:N ratio, organic matter, elevation, and mean annual precipitation (Table 2). Although these results followed the same trend in ITS1 and 18S data, however, their relative rankings varied. For example, fungal richness from ITS1 data was most strongly correlated with bulk density and organic matter, while richness from 18S data was more strongly correlated to C:N ratio and elevation in addition to bulk density (Table 2). Furthermore, there were some relationships unique to each dataset. Significant negative relationships were observed between richness and soil water repellency. Similarly, richness derived from 18S data was negatively related to total C and sand content of soil but also positively related to clay content.

We found pH was the best predictor of  $\beta$ -diversity from linear fitting for fungi no matter what gene region is amplified (Table 3 and Table 4). All fitted variables were significantly correlated to  $\beta$ -diversity, though most of these only weakly. It is likely that they did not strongly influence the fungal communities. Variables followed similar rankings in both the ITS1 and 18S data. Elevation, annual precipitation, soil moisture, C:N ratio, organic matter, and bulk density all had  $R^2$  values greater than 0.35, but their relative order differed between datasets (Table 3 and Table 4).

#### *Effect of Land Use on Functional Diversity*

There was a distinct difference in trophic modes of OTUs that were successfully matched to the FUNGUILD database between ITS1 and 18S datasets. In total, 3,402 and 1,783 OTUs from the ITS1 and 18S datasets respectively were matched to the

FUNGUILD database. Overall, saprotrophs were the most abundant trophic mode in both datasets (Fig. 6); however, pathotrophs ranked second in ITS1 (Fig. 6A) data while the pathotroph-saprotroph-symbiotroph multi-trophic group was second-most abundant in 18S data (Fig. 6B). Across land uses, proportions of pathotrophs and pathotroph-saprotroph-symbiotrophs fell with declining productivity (Fig. 7). In matches from the ITS1 data, pathotroph-saprotrophs increased across the productivity gradient (Fig. 7A), as did saprotrophs in the 18S data (Fig. 7B). The aforementioned trend in proportional abundance of pathotrophs and pathotroph-saprotroph-symbiotrophs was also present across organic matter classes (Fig. S8). Symbiotrophs appeared to follow an opposite trend, increasing as productivity fell. Interestingly, this was the case for saprotrophs in the 18S (Fig. S8B) but not the ITS1 (Fig. S8A) dataset. Proportional abundances of fungal OTUs grouped by trophic modes did not follow a discernable pattern across changing soil types (Fig. S9). For simplicity, we focused further analyses only on the broadly defined saprotroph, pathotroph, and symbiotroph groups, ignoring all combination groups; pairwise differences for all of the following comparisons are described in the Supplementary Material.

Across land uses, significant differences were observed in the richness of saprotrophic fungi in both the ITS1 ( $F_{6,258} = 25.14$ ,  $p < 0.001$ ) and 18S ( $F_{6,267} = 31.10$ ,  $p < 0.001$ ) data; however, there were differences between datasets (Fig. 8). In the ITS1 dataset, richness followed the same trend as overall fungal richness, with the highest and lowest values in the Crops/weeds and Heath/bog AVCs respectively (Fig. 8A). Although this pattern was preserved in the 18S data (Fig. 8B), richness of saprotrophs was much more even across AVCs in this case. Indeed, rather than the linear decline of richness

415 along the productivity gradient, there appeared to be 3 distinct levels in the data affiliated  
416 with (i) grassland/agricultural sites, (ii) woodlands, and (iii) bogs.

417 The same pattern was also apparent across organic matter classifications in both  
418 datasets (ITS1:  $F_{3, 260} = 32.86$ ,  $p < 0.001$ ; 18S:  $F_{3, 269} = 41.13$ ,  $p < 0.001$ ; Fig. 9). In the  
419 ITS1 dataset, each class was significantly different from the others (Fig. 9A). In the 18S  
420 data, saprotroph richness was significantly higher in mineral and humus-mineral soils  
421 than organo-mineral and organic soils (all  $p < 0.001$  except mineral – organo-mineral  $p =$   
422  $0.02$ ) (Fig. 9B). Again, the overarching trend of fungal richness was not apparent when  
423 samples were grouped by soil type. Although there were significant differences across  
424 soil types in both the ITS1 ( $F_{5, 259} = 9.7$ ,  $p < 0.001$ ) and 18S ( $F_{5, 268} = 10.73$ ,  $p < 0.001$ )  
425 datasets, these differences did demonstrate consistent patterns across soil types (Fig.  
426 S10).

427 In the case of pathotrophic fungi, richness also followed a similar trend to the  
428 saprotrophs across both datasets. In the ITS1 data, significantly ( $F_{6, 258} = 26.11$ ,  $p <$   
429  $0.001$ ) greater richness values were observed in Crops/weeds and grassland samples (Fig.  
430 8A). Richness of pathotrophs was significantly highest in Crops/weeds sites. Again, this  
431 trend was present, though not as clear, in the 18S dataset (Fig. 8B). Significant  
432 differences ( $F_{6, 267} = 52.26$ ,  $p < 0.001$ ) were observed between AVCs, with the highest  
433 richness of pathotrophs occurring in the Fertile grassland and Crop/weeds land uses.

434 Across organic matter classes, significant differences were also observed in  
435 pathotroph richness in the ITS1 ( $F_{3, 250} = 24.91$ ,  $p < 0.001$ ) and 18S ( $F_{3, 269} = 30.49$ ,  $p <$   
436  $0.001$ ) datasets. However, in this case the trends were more apparent in the 18S data than  
437 the ITS1 data (Fig. 9). Pathotroph richness was highest in mineral soils and lowest in

organic soils when compared to all other classes in the ITS1 data (Fig. 9A). However, all organic matter classifications were statistically different from each other in the 18S data (Fig. 9B), in descending order from mineral to peat soils. Again, trends were less clear across soil types (Fig. S10). Significant differences were observed in the ITS1 data ( $F_{5, 259} = 6.93$ ,  $p < 0.001$ ) with the lowest pathotroph richness found in peat soils (Fig. S10A). In the 18S data, differences between pathotrophic fungi across soil types were more similar to those observed in other groups (Fig. S10B). Pathotroph richness was significantly ( $F_{5, 268} = 13.6$ ,  $p < 0.001$ ) different across soil types with the highest values found in brown soils and the lowest in peats.

The previously described trend of declining richness across the land use productivity gradient (i.e. Fig. 4) was not apparent when considering symbiotrophs. Furthermore, although significant differences were apparent in both the ITS1 ( $F_{6, 258} = 14.88$ ,  $p < 0.001$ ) and 18S ( $F_{6, 267} = 55.13$ ,  $p < 0.001$ ) datasets they were by no means identical (Fig. 8). Symbiotroph richness was highest in Lowland wood sites followed by Upland wood. This trend was not apparent in the 18S dataset, however (Fig. 8B). Here richness of symbiotrophs was greatest in grassland AVCs and lowest in Heath/bog sites much like the overarching trend of total fungal OTU richness.

When samples were grouped by organic matter class, further discrepancies became apparent between the datasets. Whereas the previously described trend of decreasing richness with increasing organic matter content held true in the 18S data ( $F_{3, 269} = 36.28$ ,  $p < 0.001$ ; Fig. 9B), no significant differences were observed in the ITS1 dataset ( $F_{3, 260} = 1.88$ ,  $p = 0.13$ ; Fig 9A). In the 18S data, richness of symbiotrophs was greater in mineral and humus-mineral soils when compared to organo-mineral ( $p = 0.002$ ,

p = 0.04, respectively) and organic ( $p < 0.001$ ) soils (Fig. 9B). There were also no significant differences ( $F_{5, 259} = 1.43$ ,  $p = 0.21$ ) in symbiotroph richness across soil types in ITS1 data (Fig. S10A), though there were in 18S data ( $F_{5, 259} = 12.52$ ,  $p < 0.001$ ; Fig. S10B). As described previously, richness was lowest in peat soils and highest in brown soils.

We suspected that the differences in functional diversity observed between datasets might be a result of differential coverage of important groups. We were able to confirm this when we analysed the richness of OTUs identified as Glomeromycetes present in the 18S dataset (Fig. 10). All of the 162 Glomeromycetes OTUs were assigned as highly-probable symbiotrophs through FUNGUILD. Across land uses, richness of Glomeromycetes followed similar trends to those of symbiotrophs and saprotrophs from 18S data. There were significant ( $F_{6, 244} = 33.47$ ,  $p < 0.001$ ) differences across land uses, though they appeared, like the saprotroph richness to be tiered between grasslands, woods, and bogs (Fig. 10A). Richness of Glomeromycetes was higher in grasslands than all other AVCs except Crops/weeds and lowest in Heath/bog sites. Again, when grouped by organic matter class (Fig. 10B) and soil type (Fig. 10C), Glomeromycetes richness followed the same trend as saprotrophs and symbiotrophs from the 18S dataset. Richness was significantly ( $F_{3, 246} = 37.65$ ,  $p < 0.001$ ) greater in mineral and humus-mineral soils than all others. Across soil types, richness of Glomeromycetes was significantly ( $F_{5, 245} = 8.65$ ,  $p < 0.001$ ) lower in peat soils when compared to most other soil types.

#### *Relationships Between Soil Properties and Fungal Functional Diversity*

Across all samples, PLS and VIP analyses highlighted strong correlations between fungal richness and soil properties by trophic groups. Richness of pathotrophs showed similar relationships to soil properties in both datasets. There were significant, positive relationships of richness with pH and bulk density; and significant negative correlations between richness and total C, C:N ratio, organic matter, elevation, and mean annual precipitation (Table 5). As with the total fungal data, the relative rankings of the strength of relationships between pathotroph and each property varied between datasets. Organic matter was most strongly correlated with pathotroph richness from ITS1 data whereas pH was most strongly correlated with pathotroph richness in the 18S data (Table 5). Also soil moisture content was also negatively correlated with pathotroph richness in the ITS1 dataset only.

Organic matter, elevation (both negative), pH, and bulk density (both positive) all showed significant relationships with saprotroph richness in both datasets (Table 5). The correlations between richness of saprotrophs and both bulk density and pH were the strongest observed in the ITS1 data. There were also negative correlations between saprotroph richness and total C, mean annual precipitation, soil moisture, soil water repellency, and mite abundance in the ITS1 data. However, it again should be noted that the correlation with mites was extremely weak. C:N ratio was strongly and positively correlated with saprotroph richness in the 18S data. Similarly, richness from 18S data was negatively related to total C and sand content of soil but also positively related to clay content. In addition, there was a significant, positive, but weak correlation between sand content and saprotroph richness.

In both datasets, symbiotroph richness was significantly correlated with pH and C:N ratio (Table 5). Interestingly, the relationships were positive in the case of C:N ratio and negative for pH in ITS1 data but the opposite was apparent in the 18S data. There were also many more relationships unique to each dataset. Weak but significant positive relationships were observed between symbiotroph richness and rock volume, Collembola abundance, and temperature as well as a negative correlation to soil moisture. In the 18S data, stronger relationships were observed between symbiotroph richness and bulk density (positive) and elevation (negative). Furthermore a weakly negative correlation was observed with sand content in addition to weak positive correlations with clay content and total P.

## DISCUSSION

### *Primer Choice and the Total Fungal Community*

We observed congruent patterns in total fungal OTU richness across land uses, organic matter classes and soil type when measured with either ITS1 or 18S primer sets. Richness was greater in arable and grassland land uses, which are highly productive, intensively managed and declined in the less productive, largely unmanaged bogs. Although these findings had been previously known from the ITS1 dataset (George et al., 2019), it is important to note that the trend was also present in the fungal OTUs identified from 18S sequencing. A similar trend was observed across organic matter classes. Here, fungal richness fell as organic matter increased. Fungal  $\alpha$ -diversity is known to be greater in arable soils than in grasslands or forests (Szoboszlay et al., 2016). Potential mechanisms for this include: (i) increased nutrient availability due to fertiliser input



(Szoboszlay et al., 2016), and (ii) beneficial disturbance from tillage and other standard agricultural practices. The latter is consistent with the intermediate disturbance hypothesis whereby high levels of diversity are maintained by consistent interruption of successional processes (Connell, 1978).

Soils rich in organic matter, especially peats, found in upland moors, bogs, and other wetlands across harbour distinct fungal communities from neighbouring habitats (Anderson et al., 2003). Fungi dominate microbial communities in bogs (Thormann and Rice, 2007) although their proportional abundance drops sharply below the first 5 cm of bog habitats (Potter et al., 2017). Yet, richness in bogs is consistently low, perhaps due to environmental pressures such as high acidity, highly recalcitrant SOM, low nutrients and oxygen levels (Rousk et al., 2010; Tedersoo et al., 2010) or reduced competition within the fungal community.

In comparison to AVC and SOM levels, differences in fungal communities were not as clear across soil types as defined by the National Soil Map (Avery, 1980), which is inline with previous work on microbial activity across the UK (Jones et al., 2014). Richness was highest in brown soils and was lowest in peats. Brown soils commonly support grassland communities across Wales (Avery, 1980; Rudeforth et al., 1984). Nearly half of the Fertile and Infertile grasslands surveyed in GMEP were categorised as brown soils. The absence of other major trends besides these may be due to the use of the dominant soil type and lack of resolution for the soil classification. The soils map used in this study simply does not provide enough resolution (1:63, 360; Avery, 1980) for soil type to be an effective category. Furthermore, this system heavily uses subsoil properties to determine soil type (Avery, 1980), while our work only involved the upper 15 cm.

551 However, it is our opinion that the use of organic matter classification is more effective  
552 and simple metric that can be easily implemented in large-scale studies in lieu of fine-  
553 scale maps.

554 Results of PLS analyses demonstrates that soil properties and associated  
555 environmental factors influencing fungal richness are consistent across ITS1 and 18S  
556 datasets. Major drivers included pH, bulk density, C:N ratio, organic matter, elevation,  
557 and mean annual temperature (Table 2). Such results from 18S data are consistent with  
558 previous findings from the ITS1 data (George et al., 2019). However, there were certain  
559 properties that were significant in only one of the datasets and the relative importance of  
560 these properties does vary between the two datasets. There are several possible  
561 explanations for this. Firstly, 9 more samples were used in the 18S dataset ( $n = 422$ ) than  
562 the ITS1 data ( $n = 413$ ), which may have introduced the discrepancy in relative  
563 importance of the data. However, it is much more likely that a differential coverage of  
564 fungal groups between the two datasets caused these discrepancies.

565 Community composition showed consistent clustering across land uses, organic  
566 matter classes, and soil types in both data sets. As in George et al. (2019), communities  
567 were most similar in the grassland and arable sites and more spread out across woodlands  
568 and upland habitats. This was likely driven by environmental factors across Wales. In  
569 both datasets, pH was the most important environmental variable influencing community  
570 composition and although the remaining properties followed similar patterns, their  
571 relative importance again differed in the dataset. The importance of pH, elevation, C:N  
572 ratio, and precipitation in determining fungal community composition fits well in the  
573 wider context of soil fungi biogeography. Tederloo et al. (2014) previously highlighted

the importance of these variables in the distribution of fungi at the global scale. Furthermore, the strong positive correlation with C:N ratio is indicative of the expected fungal dominance (de Vries et al., 2006) of nutrient-poor, acidic soils (Bloem et al., 1997).

#### *Primer Choice and Fungal Functional Diversity*

Differences between richness of trophic modes of fungi, used here as a proxy for functional diversity, showed some discrepancies across land uses and soil classification between data sets. Saprotrophs made up the largest proportion of the 3 functional groups studied and generally exhibited the same trends as total richness across soils and land uses. This was also the case for pathotrophs. Indeed, correlations between environmental variables with pathotroph and saprotroph richness were largely consistent across datasets. However, we observed divergent trends in symbiotroph richness across land uses and soils. Symbiotroph richness was highest in woodlands in the ITS1 dataset whereas it was highest in grasslands according to the 18S data (Fig. 7A and 7B). A similar increase in richness within grasslands in the 18S data is repeated when Glomeromycetes were considered on their own (Fig. 9); AMF are the predominant mycorrhizal fungi in grassland systems (Smith and Read, 2008). The symbiotroph peak in the ITS1 data may be explained by an increase in coverage of ectomycorrhizas which are the most common group to associate with trees and shrubs (Smith and Read, 2008). Despite these differences, both datasets suggest that symbiotroph richness was low in arable land, which is in line with previous findings demonstrating high susceptibility of mycorrhizal fungi to disturbance, for example tillage (Schnoor et al., 2011; Säle et al., 2015), and the

addition of fertilizers, which decreases the receptiveness of many agricultural plants to mycorrhizal infection (Smith and Read, 2008).

The divergent trend in symbiotroph richness and discrepancies in relationships between functional groups and environmental variables likely stem from primer biases. Primer biases have been well recognised as a confounding factor in categorising communities from environmental DNA (Cai et al., 2013; Elbrecht and Leese, 2015; Tedersoo et al., 2015). Tedersoo et al. (2015) assessed the effectiveness of fungal barcodes from the ITS, 18S, and 28S rDNA regions and found that primer choice did not affect richness or  $\beta$ -diversity results of soil fungi communities from Papua New Guinea, although fewer OTUs were recovered by 18S primers than ITS primers. *In silico* analyses suggests such findings are the result of lumping of sequences in the 18S that may predominantly affect rare sequences, thereby strengthening community matrices. Similarly, results were similar enough for all primers to be suitable for analyses at the class-level (Tedersoo et al., 2015). Although the 18S primers used here were designed to cover the breadth of eukaryotes and may lack specificity to fungi (Behnke et al., 2011), our results show strong congruence to the ITS1 data across total richness and indeed most functional groups.

Unlike Tedersoo et al. (2015) we observed considerable differences in the proportions of fungal classes between the ITS1 and 18S data sets. We suspect that such differences stem from the need to use appropriate databases to assign taxonomy to OTUs to each dataset (Xue et al., 2019). Perhaps only 30%-35% of Glomeromycetes are present in 18S and ITS databases, respectively (Hart et al., 2015), and although sequences are continuously being uploaded to such repositories, it is likely the majority of AMF are not

identifiable from environmental samples (but see Öpik et al., 2014). Similarly we suspect that, although not studied in detail, primer choice may lead to biases in other groups. Archaeorhizomycetes accounted for nearly 25% of the 18S sequences but less than 1% from the ITS1 data (Fig. 1B). Primer bias has been recognised for Archaeorhizomycetes even before the class' formal description; approximately 19% of 18S sequences collected from Anderson et al. (2003), have been matched to Archaeorhizomycetes, whereas none were recovered from the same samples using ITS primers. Despite its recent description, Archaeorhizomycetes are ubiquitous components of soil communities. Strong associations have been observed with trees, yet precise functional roles of these fungi have yet to be determined (Rosling et al., 2011). Subsequently, such biases likely account for divergent relationships between functional group richness and environmental properties.

### *Conclusions*

Our comparison of the use of ITS1 and 18S primers and their respective databases in a nationwide metabarcoding survey of fungi yielded 3 major findings. First, the congruent findings of total richness and  $\beta$ -diversity across land use and their relationships to environmental variables confirmed our previous research (George et al., 2019). Second, soil organic matter was found to be a more sensitive metric than soil type in our survey design. Third, biases from the combination of primer and database choice became apparent for certain classes of fungi, including Glomeromycetes and Archaeorhizomycetes, which strongly influenced functional group richness across land uses as well as their relationships with environmental variables. It is therefore important

to recognise the sensitivity of metabarcoding to primer choice, even when using universal primers. Without simultaneous analyses of environmental DNA using both primers and databases, the presence of AM fungi as well as the newly characterised Archaeorhizomycetes would have been overlooked and unquantified in this survey. Furthermore, since the majority of soil biodiversity is undescribed (Ramirez et al., 2015), utilising multiple primers will elucidate a more complete picture of belowground biodiversity by revealing shortcomings in existing probes and revealing the presence of as yet undescribed organisms. We therefore advocate that future nation-wide surveys included both a sample-based metric of soil type (i.e. organic matter classification) and multiple primers for fungal biodiversity. Such measures should not be arduous to implement, especially if researchers can identify specific fungal groups of particular interest to accommodate.

## **AUTHOR CONTRIBUTIONS**

P.B.L.G., D.L.J., D.A.R. and S.C. conceived this project. Bioinformatics and statistical analyses were led by P.B.L.G. with assistance from S.C. and R.I.G. P.B.L.G. wrote the first draft of the manuscript and S.C., D.A.R., and D.L.J. contributed to subsequent revisions. All authors read and approved the final draft of the manuscript.

## **FUNDING**

This project was supported as part of the Welsh Government funded Glastir Monitoring & Evaluation Programme (Contract reference: C147/2010/11) and by the UK Natural Environment Research Council (NERC) through the RC/Centre for Ecology &

Hydrology (CEH Project: NEC04782). PBLG was supported by a Soils Training and Research Studentship (STARS) grant from the Biotechnology and Biological Sciences Research Council and NERC [Grant number NE/M009106/1]. STARS is a consortium consisting of Bangor University, British Geological Survey, Centre for Ecology and Hydrology, Cranfield University, James Hutton Institute, Lancaster University, Rothamsted Research and the University of Nottingham.

## **ACKNOWLEDGEMENTS**

This research was supported under the Glastir Monitoring & Evaluation Programme (Contract reference: C147/2010/11) and by the UK Natural Environment Research Council (NERC) through the RC/Centre for Ecology & Hydrology (CEH Project: NEC04782). We thank the GMEP Team for their contribution in collecting the data and the laboratory staff of Environment Centre Wales, especially Delphine Lallias, for sample processing and DNA extraction. Thanks also to Fiona M. Seaton for assistance with statistical analyses. We thank Supercomputing Wales for support and training on their system for bioinformatics analyses. We thank John G. Kenny and Richard M. Eccles for generating sequence data at the Centre for Genomic Research, University of Liverpool.

## **SUPPLEMENTARY MATERIAL**

The Supplementary Material for this article can be access online at: []

## **DATA AVAILABILITY**

Data associated with this paper will be publically published in the National Environment Research Council (NERC) Environmental Information Data Centre (EIDC). Data are also available from the authors upon reasonable request with permission from the Welsh Government. Sequences with limited sample metadata have been uploaded to the

European Nucleotide Archive and can be accessed with the following primary accession codes after the end of data embargo: PRJEB28028 (ITS1), and PRJEB28067 (18S).

## REFERENCES

- Altschul, S. F., Gish, W., Miller, W., Myers, E. W., and Lipman, D. J. (1990) Basic local alignment search tool. *J. Mol. Biol.* 215, 403-410. doi: 10.1016/S0022-2836(05)80360-2
- Anderson, I. C., Campbell, C. D., and Prosser, J. I. (2003). Diversity of fungi in organic soils under a moorland – Scots pine (*Pinus sylvestris* L). *Environ. Microbiol.* 5, 1121-1132. doi: 10.1046/j.1462-2920.2003.005522.x
- Avery, .W. (1980). Soil classifications for England and Wales (Higher Categories). Soil Survey Technical Monograph No 14. Harpenden, UK.
- Behnke, A., Engel, M., Christen, R., Nebel, M., Klein, R. R., and Stoeck, T. (2011). Depicting more accurate pictures of protistan community complexity using pyrosequencing of hypervariable SSU rRNA gene regions. *Environ. Microbiol.* 13, 340-349. doi: 10.1111/j.1462-2920.2010.02332.x
- Berruti, A., Desirò, A., Visentin, S., Zecca, O., Bonfante, P. (2017). ITS fungal barcoding primers versus 18S AMF-specific primers reveal similar AMF-based diversity patterns in roots and soils of three mountain vineyards. *Environ. Microbiol. Rep.* 9, 658-667.
- Bloem, J., de Ruiter, P. C., and Bouwman, L. A. (1997). “Soil food webs and nutrient cycling in agroecosystems,” in *Modern soil microbiology*, eds. van Elsas, J. D., Jansson, J. K., and Trevors, J. T. (Boca Raton, FL: CRC Press), 245–278.
- Bunce, R. G. H., Carr, C. J., Gillespie, M. K., Howard, D. C., Scott, W. A., Smart, S. M., Van de Poll, H. M., and Watkins, J. W. (1999). Vegetation of the British countryside – the countryside vegetation system. (Department of the Environment, Transport and the Regions, London).
- Cai, L., Ye, L., Tong, A. H. Y., Lok, S., and Zhang, T. (2013). Biased diversity metrics revealed by bacterial 16S phylotags derived from different primer sets. *PLoS ONE* 8, e53649. doi: 10.1371/journal.pone.0053649
- Caporaso, J. G., Kuczynski, J., Stombaugh, J., Bittinger, K., Bushman, F. D., Costello, E. K., Fierer, N., Peña, A. G., Goodrich, J. K., Gordon, J. I., Huttley, G. A., Kelley, S. T., Knights, D., Koenig, J. E., Ley, R. E., Lozupone, C. A., McDonald, D., Muegge, B. D., Pirrung, M., Reeder, J., Sevinsky, J. R., Turnbaugh, P. J., Walters, W. A., Wildmann, J., Yaysunen, T., Zaneveld, J., and Knight, R. (2010). QIIME allows analysis of high-throughput community sequencing data. *Nat. Methods* 7, 335-336. doi: 10.1038/nmeth.f.303
- Cavender-Bares, J., Kozak, K. H., Fine, P. V. A., Kembel, S. W. (2009) The merging of community ecology and phylogenetic biology. *Ecol. Lett.* 12, 693-715. doi: 10.1111/j.1461-02548.2009.01314.x
- Connell, J. H. (1978). Diversity in tropical rain forests and coral reefs. *Science* **199**, 1302-1310. doi: 10.1126/science.199.4335.1302
- Chong, I.-G., and Jun, C.-H. (2005). Performance of some variable selection methods when multicollinearity is present. *Chemometr. Intell. Lab. Syst.* 78, 103-112. doi: 10.1016/j.chemolab.2004.12.011



- Deshpande, V., Wang, Q., Greenfield, P., Charleston, M., Porras-Alfaro, A., Kuske, C. R., Cole, J. R., Midgley, D. J., and Tran-Dinh, N. (2016). Fungal identification using a Bayesian classifier and the Warcup training set of internal transcribed spacer sequences. *Mycologia* 108, 1-5. doi: 10.3852/14-293.
- Edgar, R. C. (2010). Search and clustering orders of magnitude faster than BLAST. *Bioinformatics* 10, 2460-2461. doi: 10.1093/bioinformatics/btq461
- Elbrecht, V., and Leese, F. (2015). Can DNA-based ecosystem assessments quantify species abundance? Testing primer bias and biomass-sequence relationships with an innovative metabarcoding protocol. *PLoS ONE* 10, e0130324. doi: 10.1371/journal.pone.0130324
- Emmett, B. A. & the GMEP Team. (2017). Glastir Monitoring & Evaluation Programme. Final Report to Welsh Government (Contract reference: C147/2010/11. NERC/Centre for Ecology & Hydrology (CEH Project: NEC04780/NEC05371/NEC05782).
- Emmett, B. A., Reynolds, B., Chamberlain, P. M., Rowe, E., Spurgeon, D., Brittain, S. A., Frogbrook, Z., Hughes, S., Lawlor, A. J., Poskitt, J., Potter, E., Robinson, D. A., Scott, A., Wood, C., and Woods, C. (2010). Countryside Survey: Soils Report from 2007. NERC/Centre for Ecology and Hydrology (CS Technical Report No. 9/07, CEH Project Number: C03259).
- Epp, L. S., Boessenkool, S., Bellemain, E. P., Haile, J., Esposito, A., Riaz, T., Erséus, C., Gusarov, V. I., Edwards, M. E., Johnsen, A., Stenøien, H. K., Hassel, K., Kauserud, H., Yoccoz, N. G., Bråthen, K. A., Willerslev, E., Taberlet, P., Coissac, E., and Brochmann, C. (2012). New environmental metabarcodes for analysing soil DNA: potential for studying past and present ecosystems. *Mol. Ecol.* 8, 1821-1833. doi: 10.1111/j.1365-294X.2012.05537.x
- Fisher, M. C., Henk, D. A., Briggs, C. J., Brownstein, J. S., Madoff, L. C., McCraw, S. L., and Gurr, S. J. (2012). Emerging fungal threats to animal, plant and ecosystem health. *Nature* 484, 186-194. doi: 10.1038/nature10947
- George, P. B. L., Lallias, D., Creer, S., Seaton, F. M., Kenny, J. G., Eccles, R. M., Griffiths, R. I., Lebron, I., Emmett, B. A., Robinson, D. A., and Jones, D. L. (2019). Divergent national-scale trends of microbial and animal biodiversity revealed across diverse temperate soil ecosystems. *Nat. Commun.* 10, 1107. doi:10.1038/s41467-019-09031-1
- Hart, M. M., Aleklett, K., Chagnon, P.-L., Egan, C., Ghignone, S., Helgasson, T., Lekberg, Y., Öpik, M., Pickles, B. J., and Waller, L. (2015). Navigating the labyrinth: a guide to sequence-based, community ecology of arbuscular mycorrhizal fungi. *New Phytol.* 207, 235-247. doi: 10.1111/nph.13340
- Hothorn, T., Bretz, F., and Westfall, P. (2008). Simultaneous inference in general Parametric Models. *Biometrical J.* 50, 346-363. doi: 10.1002/bimj.200810425
- Jones, D. L., Simfukwe, P., Hill, P. W., Mills, R. T. E., Emmett, B. A. Evaluation of dissolved organic carbon as a soil quality indicator in national monitoring schemes. *PLoS One* 9(3):e90882. doi: 10.1371/journal.pone.0090882
- Kõljalg, U., Nilsson, R. H., Abarenkov, K., Tedersoo, L., Taylor, A. F. S., Bahram, M., Bates, S. T., Burns, T. D., Bengtsson-Palme, J., Callaghan, T. M., Douglas, B., Drenkhan, T., Eberhardt, U., Dueñas, M., Grebenc, T., Griffith, G. W., Hartmann, M., Kirk, P. M., Kohout, P., Larsson, E., Lindahl, B. D., Lücking, R., Martín, M. P., Matheny, P. B., Nguyen, N. H., Niskanen, T., Oja, J., Peay, K. G., Peintner, U.,

- Peterson, M., Põldmaa, K., Saag, L., Saar, I., Senés, C., Smith, M. E., Suija, A., Taylor, D. L., Telleria, M. T., Weiss, M., and Larsson, K.-H. (2013). Towards a unified paradigm for sequence-based identification of fungi. *Mol. Ecol.* 22, 5271-5277. doi: 10.1111/mec.12481
- Lallias, D., Hiddink, J., G., Fonseca, V. G., Gaspar, J. M., Sung, W., Neill, S. P., Barnes, N., Ferrero, T., Hall, N., Lambshead, P. J. D., Packer, M., Thomas, W. K., and Creer, S. (2015). Environmental metabarcoding reveals heterogeneous drivers of microbial eukaryote diversity in contrasting estuarine ecosystems. *ISME J.* 9, 1208-1221. doi: 10.1038/ismej.2014.213
- Martin, M. (2011). Cutadapt removes adapter sequences from high-throughput sequencing reads. *EMBnet.journal*. 2011. Date of access 04/07/2017. *EMBnet* 17, 10-12. doi: 10.14806/ej.17.1.200
- McMurdie, P. J., and Holmes, S. (2013). phyloseq: an R package for reproducible interactive analysis and graphics of microbiome census data. *PLoS ONE* 8, e61217. doi: 10.1371/journal.pone.0061217
- Mevik, B.-H., Wehrens, R., Liland, K. H., and Hiemstra, P. (2016). pls: Partial Least Squares and Principal Component Regression. R package version 2.6-0. <https://CRAN.R-project.org/package=pls>
- Nguyen, N. H., Song, Z., Bates, S. T., Branco, S., Tedersoo, L., Menke, J., Schilling, J. S., and Kennedy, P. G. (2016). FUNGuild: an open annotation tool for parsing fungal community datasets by ecological guild. *Fungal Ecol.* 20, 241-248. doi: 10.1016/j.funeco.2015.06.006
- Oksanen, J., Blanchet, F. G., Friendly, M., Kindt, R., Legendre, P., McGlinn, D., Minchin, P. R., O'Hara, R. B., Simpson, G. L., Solymos, P., Henry, M., Stevens, H., Szoecs, E., and Wagner, H. (2016). Vegan: community ecology package. R package version 2.4-0. <https://CRAN.R-project.org/package=vegan>.
- Öpik, M., Davison, J., Moora, M., and Zobel, M. (2014). DNA-based detection and identification of Glomeromycota: the virtual taxonomy of environmental sequences<sup>1</sup>. *Botany* 92, 135-147. doi: 10.1139/cjb-2013-0110
- Peay, K. G., Kennedy, G., and Bruns, T. D. (2008). Fungal community ecology: a hybrid beast with a molecular master. *BioScience* 58, 799-810. doi: 10.1641/B580907
- Pinheiro, J., Bates, D., DebRoy, S., Sarkar, D., EISPAK authors, Heisterkamp, S., Van Willigen, B., and R Core Team (2016). nlme: Linear and Nonlinear Mixed Effects Models. R package version 3.1e138. <https://CRAN.R-project.org/package=nlme>
- Potter, C., Freeman, C., Golyshin, P. N., Ackermann, G., Fenner, N., McDonald, J. E., Ehbair, A., Jones, T. G., Murphy, L. M., and Creer, S. (2017). Subtle shifts in microbial communities occur alongside the release of carbon induced by drought and rewetting in contrasting peatland ecosystems. *Sci. Rep.-UK* 7, 11314. doi: 10.1038/s41598-017-11546
- Quast, C., Pruesse, E., Yilmaz, P., Gerken, J., Schweer, T., Yarza, P., Peplies, J., and Glöckner, F. O. (2013). The SILVA ribosomal RNA gene database project: improved data processing and web-based tools. *Nucl. Acids Res.* 41 (D1), D590-D596. doi: 10.1093/nar/gks1219
- R Core Team (2017). R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria (2017). <https://www.R-project.org/>.

- Ramirez, K. S., Leff, J. W., Barberán, A., Bates, S. T., Betley, J., Crowther, T. W., Kelly, E. F., Oldfield, E. E., Shaw, E. A., Steenbock, C., Bradford, M. A., Wall, D. H., Fierer, N. (2015). Biogeographic patterns in below-ground diversity in New York City's Central Park are similar to those observed globally. *Proc. R. Soc. B.* 281, 20141988. doi: 10.1098/rspb.2014.1988
- Robinson, E. L., Blyth, E., Clark, D. B., Comyn-Platt, E., Finch, J., and Rudd, A. C. (2017). Climate hydrology and ecology research support system meteorology dataset for Great Britain (1961-2015) [CHESS-met] v1.2. NERC Environmental Information Data Centre.
- Rognes, T., Flouri, T., Nichols, B., Quince, C., and Mahé, F. (2016). VSEARCH: a versatile open source tool for metagenomics. *PeerJ* 4, e2584. doi: 10.7717/peerj.2584
- Rosling, A., Cox, F., Cruz-Martinez, K., Ihrmark, K., Grelet, G.-A., Lindahl, B. D., Menkis, A., and James, T. Y. (2011). Archaeorhizomycetes: unearthing an ancient class of ubiquitous soil fungi. *Science* 333, 876-879. doi: 10.1126/science.1206958
- Rousk, J., Bååth, E., Brookes, P. C., Lauber, C. L., Lozupone, C., Caporaso, J. G., Knight, R., and Fierer, N. (2010). Soil bacterial and fungal communities across a pH gradient in an arable soil. *ISME J.* 4, 1340-1351. doi: 10.1038/ismej.2010.58
- Rudeforth, C. C., Hartnup, R., Lea, J. W., Thompson, T. R. E., and Wright, P. S. (1984). Soils and their use in Wales. Harpenden: Lawes Agricultural Trust (Soil Survey of England and Wales).
- Sagova-Mareckova, M., Cermak, L., Novotna, J., Plhachova, K., Forstova, J., and Kopecky, J. (2008). Innovative methods for soil DNA purification tested in soils with widely differing characteristics. *Appl. Environ. Microb.* 74, 2902-2907. doi: 10.1128/AEM.02161-07
- Säle, V., Aguilera, P., Laczko, E., Mäder, P., Berner, A., Zihlmann, U., van der Heijden, M. G. A., and Oehl, F. (2015). Impact of conservation tillage and organic farming on the diversity of arbuscular mycorrhizal fungi. *Soil Biol. Biochem.* 84, 38-52. doi:10.1016/j.soilbio.2015.02.005
- Schoch, C. L., Seifert, K. A., Huhndorf, S., Robert, V., Spouge, J. L., Levesque, C. A., Chen, W., and the Fungal Barcoding Consortium, (2012). Nuclear ribosomal internal transcribed spacer (ITS) region as a universal DNA barcode marker for Fungi. *P. Natl. Acad. Sci. USA.* 109, 6241-6246. doi: 10.1073/pnas.1117018109
- Schnoor, T. K., Lekberg, Y., Rosendahl, S., and Olsson, P. A. (2011) Mechanical soil disturbance as a determinant of arbuscular mycorrhizal fungal communities in semi-natural grassland. *Mycorrhiza* 21, 211-220. doi: 10.1007/s00572-010-0325-3
- Schüßler, A., Schwarzott, D., and Walker, C. (2001). A new fungal phylum, the *Glomeromycota*: phylogeny and evolution. *Mycol. Res.* 105, 1413-1421. doi: 10.1017/S0953756201005196
- Smith, S. E., and Read, D. J. (2008). Mycorrhizal Symbiosis 3<sup>rd</sup> ed. London: Academic Press.
- Stockinger, H., Krüger, M., and Schüßler, A. (2010). DNA barcoding of arbuscular mycorrhizal fungi. *New Phytol.* 187, 461-474. doi: 10.1111/j.1469-8137.2010.03262.x
- Szoboszlai, M., Dohrmann, A. B., Poeplau, C., Don, A., and Tebbe, C. C. (2017). Impact of land-use change and soil organic carbon quality on microbial diversity in soils across Europe. *FEMS Microbiol. Ecol.* 93, fix146. doi: 10.1093/femsec/fix146

- Tedersoo, L., Anslan, S., Bahram, M., Põlme, S., Riit, T., Liiv, I., Kõljalg, U., Kisand, V., Nilsson, R. H., Hildebrand, F., Bork, P., and Abarenkov, K. (2015). Shotgun metagenomes and multiple primer pair-barcode combinations of amplicons reveal biases in metabarcoding analyses of fungi. *MycKeys* 10, 1-43. doi: 10.3897/mycokeys.104852
- Tedersoo, L., Bahram, M., Põlme, S., Kõljalg, U., Yorou, N. S., Wijesundera, R., Ruiz, L. V., Vasco-Palacios, A. M., Thu, P. Q., Suija, A., Smith, M. E., Sharp, C., Saluveer, E., Saitta, A., Rosas, M., Riit, T., Ratkowsky, D., Pritsch, K., Põldmaa, K., Piepenbring, M., Phosri, C., Peterson, M., Parts, K., Pärtel, K., Otsing, E., Nouhra, E., Njouonkou, A. L., Nilsson, R. H., Morgado, L. N., Mayor, J., May, T. W., Majuakim, L., Lodge, D. J., Lee, S. S., Larsson, K.-H., Kohout, P., Hosaka, K., Hiiesahu, I., Henkel, T. W., Harend, H., Guo, L., Greslebin, A., Grelet, G., Geml, J., Gates, G., Dunstan, W., Dunk, C., Drenkhan, R., Dearnaley, J., de Kesel, A., Dang, T., Chen, X., Buegger, F., Brearley, F. Q., Bonito, G., Anslan, S., Abell, S., and Abarenkov, K. (2014). Global diversity and geography of soil fungi. *Science* 346(6213): 1256688. doi: 10.1126/science.1256688
- Thiéry, O., Vasar, M., Jairus, T., Davison, J., Roux, C., Kivistik, P.-A., Metspalu, A., Milani, L., Saks, Ü., Moora, M., Zobel, M., and Öpik, M. Sequence variation in nuclear ribosomal small subunit, internal transcribed spacer and large subunit regions of *Rhizophagus irregularis* and *Gigaspora margarita* is high and isolate-dependent. *Mol. Ecol.* 25, 2816-2832. doi: 10.1111/mec.13655
- Thormann, M. N., and Rice, A. V. (2007). Fungi from peatlands. *Fungal Divers.* 24, 241-299.
- de Vries, F. T., Hoffland, E., van Eekeren, N., Brussard, L., and Bloem, J. (2006). Fungal/bacterial ratios in grasslands with contrasting nitrogen management. *Soil Biol. Biochem.* 38, 2092-2103. doi: 10.1016/j.soilbio.2006.01.008
- Wang, B., and Qui, Y.-L. (2006). Phylogenetic distribution and evolution of mycorrhizas in land plants. *Mycorrhiza* 16, 299-363. doi: 10.1007/s00572-005-0033-6
- Wang, Q., Garrity, G. M., Tiedje, J. M., and Cole, J. R. (2007). Naïve Bayesian classifier for rapid assignment of rRNA sequences into the new bacterial taxonomy. *Appl. Environ. Microb.* 73, 5261-5267. doi: 10.1128/AEM.00062-07
- Weiss, S., Xu, Z. Z., Peddada, S., Amir, A., Bittinger, K., Gonzalez, A., Lozupone, C., Zaneveld, J. R., Vázquez-Baeza, Y., Birmingham, A., Hyde, E. R., and Knight, R. (2017). Normalization and microbial differential abundance strategies depend upon data characteristics. *Microbiome* 5, 1-18. doi: 10.1186/s40168-017-0237-y
- Xue, C., Hao, Y., Pu, X., Penton, C. R., Wang, Q., Zhao, M., Zhang, B., Ran, W., Huang, Q., Shen, Q., and Tiedje, J. M. (2019). Effect of LSU and ITS genetic markers and reference databases on analyses of fungal communities. *Biol. Fert. Soils* 55, 79-88. doi: 10.1007/s00374-018-1331-4
- Yuste, J. C., Peñuelas, J., Estiarte, M., Garcia-Mas, J., Mattana, S., Ogaya, R., Pujol, M., and Sardans, J. (2011). Drought-resistant fungi control soil organic matter decomposition and its response to temperature. *Glob. Change Biol.* 17, 1475-1486. doi: 10.1111/j.1365-2486.2010.02300.x

Captions

**Fig. 1.** Map of sites selected for GMEP monitoring. To protect landowner anonymity, each triangle gives an approximate location of every 1 km<sup>2</sup> plot from which samples were taken

**Fig. 2.** Composition of fungal classes from ITS1 and 18S datasets. **A)** Venn diagram denoting total number of shared and unique classes in each data set, following exclusion of unknown sequences. Sankey diagrams of proportional abundances of fungal OTUs from all samples from **B)** ITS1 data and **C)** 18S data. Arms denote proportions of OTUs of the most populous classes.

**Fig. 3.** Proportionate abundances of fungal OTUs for **A)** ITS1 and **B)** 18S data across Aggregate Vegetation Class. Aggregate Vegetation Classes are ordered from most (Crops/weeds) to least (Heath/bog) productive.

**Fig. 4.** Boxplots of fungal OTU richness for **A)** ITS1 and **B)** 18S datasets plotted against Aggregate Vegetation Class. Aggregate Vegetation Classes are ordered from most (Crops/weeds) to least (Heath/bog) productive. Boxes cover the first and third quartiles and horizontal lines denote the median. Black dots represent outliers beyond the whiskers, which cover 1.5X the interquartile range. Notches indicate confidence interval around the median. Overlapping notches are a proxy for non-significant differences between medians. Black dots are outliers.

**Fig. 5.** Boxplots of fungal OTU richness for **A)** ITS1 and **B)** 18S datasets plotted against organic matter class. Organic matter classes are listed in order of increasing percent organic matter. Boxes cover the first and third quartiles and horizontal lines denote the median. Black dots represent outliers beyond the whiskers, which cover 1.5X the interquartile range. Notches indicate confidence interval around the median. Overlapping notches are a proxy for non-significant differences between medians. Black dots are outliers.

**Fig. 6.** Non-metric dimensional scaling ordinations of fungal community composition across GMEP sites. Samples are coloured by Aggregate Vegetation Class. Data from ITS1 (stress = 0.13) is shown in **A)**; Data from 18S (stress = 0.11) is shown in **B)**.

**Fig. 7.** Proportionate abundances of fungal OTUs matched to FUNGuild trophic groups for **A)** ITS1 and **B)** 18S data across Aggregate Vegetation Classes. Aggregate Vegetation Classes are ordered from most (Crops/weeds) to least (Heath/bog) productive. Abbreviations for multi-trophic mode groups are as follows: Path.-Sap. (Pathotroph-Saprotroph); Path.-Sap.-Sym. (Pathotroph-Saprotroph-Symbiotroph); Path.-Sym. (Pathotroph-Symbiotroph); Sap.-Path.-Sym (Saprotroph-Pathotroph-Symbiotroph); Sap.-Sym. (Saprotroph-Symbiotroph).

**Fig. 8.** Boxplots of richness of fungal OTUs matched to the pathotrophic, saprotroph, and symbiotroph trophic modes in FUNGuild for **A)** ITS1 and **B)** 18S datasets plotted against Aggregate Vegetation Class. Aggregate Vegetation Classes are ordered from most

(Crops/weeds) to least (Heath/bog) productive. Boxes cover the first and third quartiles and horizontal lines denote the median. Black dots represent outliers beyond the whiskers, which cover 1.5X the interquartile range. Notches indicate confidence interval around the median. Overlapping notches are a proxy for non-significant differences between medians. Black dots are outliers.

**Fig. 9.** Boxplots of richness of fungal OTUs matched to the pathotrophic, saprotroph, and symbiotroph trophic modes in FUNGuild for **A)** ITS1 and **B)** 18S datasets plotted against organic matter class. Organic matter classes are listed in order of increasing percent organic matter. Boxes cover the first and third quartiles and horizontal lines denote the median. Black dots represent outliers beyond the whiskers, which cover 1.5X the interquartile range. Notches indicate confidence interval around the median. Overlapping notches are a proxy for non-significant differences between medians. Black dots are outliers.

**Fig. 10.** Boxplots of richness of Glomeromycetes OTUs plotted against **A)** Aggregate Vegetation Class; **B)** organic matter class; **C)** soil type. Aggregate Vegetation Classes are ordered from most (Crops/weeds) to least (Heath/bog) productive. Organic matter classes are listed in order of increasing percent organic matter. Soils are listed in increasing order of moisture retention. Boxes cover the first and third quartiles and horizontal lines denote the median. Black dots represent outliers beyond the whiskers, which cover 1.5X the interquartile range. Notches indicate confidence interval around the median. Overlapping

981 notches are a proxy for non-significant differences between medians. Black dots are  
982 outliers.

983

984



985

986

987

988

989

990

991

992

993

994

995

996

997

998

999

1000

1001

1002

1003

1004

1005

1006

1007

1008

1009

1010

1011

1012

1013

1014

1015

1016

1017

1018

1019

1020

1021

1022

1023

1024

1025

1026

1027

1028

1029

1030

1031

1032

1033

1034

1035

1036

1037

1038

1039

1040

1041

1042

1043

1044

1045

1046

1047

1048

1049

1050

1051

1052

1053

1054

1055

1056

1057

1058

1059

1060

1061

1062

1063

1064

1065

1066

1067

1068

1069

1070

1071

1072

1073

1074

1075

1076

1077

1078

1079

1080

1081

1082

1083

1084

1085

1086

1087

1088

1089

1090

1091

1092

1093

1094

1095

1096

1097

1098

1099

1100

1101

1102

1103

1104

1105

1106

1107

1108

1109

1110

1111

1112

1113

1114

1115

1116

1117

1118

1119

1120

1121

1122

1123

1124

1125

1126

1127

1128

1129

1130

1131

1132

1133

1134

1135

1136

1137

1138

1139

1140

1141

1142

1143

1144

1145

1146

1147

1148

1149

1150

1151

1152

1153

1154

1155

1156

1157

1158

1159

1160

1161

1162

1163

1164

1165

1166

1167

1168

1169

1170

1171

1172

1173

1174

1175

1176

1177

1178

1179

1180

1181

1182

1183

1184

1185

1186

1187

1188

1189

1190

1191

1192

1193

1194

1195

1196

1197

1198

1199

1200

1201

1202

1203

1204

1205

1206

1207

1208

1209

1210

1211

1212

1213

1214

1215

1216

1217

1218

1219

1220

1221

1222

1223

1224

1225

1226

1227

1228

1229

1230

1231

1232

1233

1234

1235

1236

1237

1238

1239

1240

1241

1242

1243

1244

1245

1246

1247

1248

1249

1250

1251

1252

1253

1254

1255

1256

1257

1258

1259

1260

1261

1262

1263

1264

1265

1266

1267

1268

1269

1270

1271

1272

1273

1274

1275

1276

1277

1278

1279

1280

1281

1282

1283

1284

1285

1286

1287

1288

1289

1290

1291

1292

1293

1294

1295

1296

1297

1298

1299

1300

1301

1302

1303

1304

1305

1306

1307

1308

1309

1310

1311

1312

1313

1314

1315

1316

1317

1318

1319

1320

1321

1322

1323

1324

1325

1326

1327

1328

1329

1330

1331

1332

1333

1334

1335

1336

1337

1338

1339

1340

1341

1342

1343

1344

1345

1346

1347

1348

1349

1350

1351

1352

1353

1354

1355

1356

1357

1358

1359

1360

1361

1362

1363

1364

1365

1366

1367

1368

1369

1370

1371

1372

1373

1374

1375

1376

1377

1378

1379

1380

1381

1382

1383

1384

1385

1386

1387

1388

1389

1390

1391

1392

1393

1394

1395

1396

1397

1398

1399

1400

1401

1402

1403

1404

1405

1406

1407

1408

1409

1410

1411

1412

1413

1414

1415

1416

1417

1418

1419

1420

1421

1422

1423

1424

1425

1426

1427

1428

1429

1430

1431

1432

1433

1434

1435

1436

1437

1438

1439

1440

1441

1442

1443

1444

1445

1446

1447

1448

1449

1450

1451

1452

1453

1454

1455

1456

1457

1458

1459

1460

1461

1462

1463

1464

1465

1466

1467

1468

1469

1470

1471

1472

1473

1474

1475

1476

1477

1478

1479

1480

1481

1482

1483

1484

1485

1486

1487

1488

1489

1490

1491

1492

1493

1494

1495

1496

1497

1498

1499

1500

1501

1502

1503

1504

1505

1506

1507

1508

1509

1510

1511

1512

1513

1514

1515

1516

1517

1518

1519

1520

1521

1522

1523

1524

1525

1526

1527

1528

1529

1530

1531

1532

1533

1534

1535

1536

1537

1538

1539

1540

1541

1542

1543

1544

1545

1546

1547

1548

1549

1550

1551

1552

1553

1554

1555

1556

1557

1558

1559

1560

1561

1562

1563

1564

1565

1566

1567

1568

1569

1570

1571

1572

1573

1574

1575

1576

1577

1578

1579

1580

1581

1582

1583

1584

1585

1586

1587

1588

1589

1590

1591

1592

1593

1594

1595

1596

1597

1598

1599

1600

1601

1602

1603

1604

1605

1606

1607

1608

1609

1610

1611

1612

1613

1614

1615

1616

1617

1618

1619

1620

1621

1622

1623

1624

1625

1626

1627

1628

1629

1630

1631

1632

1633

1634

1635

1636

1637

1638

1639

1640

1641

1642

1643

1644

1645

1646

1647

1648

1649

1650

1651

1652

1653

1654

1655

1656

1657

1658

1659

1660

1661

1662

1663

1664

1665

1666

1667

1668

1669

1670

1671

1672

1673

1674

1675

1676

1677

1678

1679

1680

1681

1682

1683

1684

1685

1686

1687

1688

1689

1690

1691

1692

1693

1694

1695

1696

1697

1698

1699

1700

1701

1702

1703

1704

1705

1706

1707

1708

1709

1710

1711

1712

1713

1714

1715

1716

1717

1718

1719

1720

1721

1722

1723

1724

1725

1726

1727

1728

1729

1730

1731

1732

1733

1734

1735

1736

1737

1738

1739

1740

1741

1742

1743

1744

1745

1746

1747

1748

1749

1750

1751

1752

1753

1754

1755

1756

1757

1758

1759

1760

1761

1762

1763

1764

1765

1766

1767

1768

1769

1770

1771

1772

1773

1774

1775

1776

1777

1778

1779

1780

1781

1782

1783

1784

1785

1786

1787

1788

1789

1790

1791

1792

1793

1794

1795

1796

1797

1798

1799

1800

1801

1802

1803

1804

1805

1806

1807

1808

1809

1810

1811

1812

1813

1814

1815

1816

1817

1818

1819

1820

1821

1822

1823

1824

1825

1826

1827

1828

1829

1830

1831

1832

1833

1834

1835

1836

1837

1838

1839

1840

1841

1842

1843

1844

1845

1846

1847

1848

1849

1850

1851

1852

1853

1854

1855

1856

1857

1858

1859

1860

1861

1862

1863

1864

1865

1866

1867

1868

1869

1870

1871

1872

1873

1874

1875

1876

1877

1878

1879

1880

1881

1882

1883

1884

1885

1886

1887

1888

1889

1890

1891

1892

1893

1894

1895

1896

1897

1898

1899

1900

1901

1902

1903

1904

1905

1906

1907

1908

1909

1910

1911

1912

1913

1914

1915

1916

1917

1918

1919

1920

1921

1922

1923

1924

1925

1926

1927

1928

1929

1930

1931

1932

1933

1934

1935

1936

1937

1938

1939

1940

1941

1942

1943

1944

1945

1946

1947

1948

1949

1950

1951

1952

1953

1954

1955

1956

1957

1958

1959

1960

1961

1962

1963

1964

1965

1966

1967

1968

1969

1970

1971

1972

1973

1974

1975

1976

1977

1978

1979

1980

1981

1982

1983

1984

1985

1986

1987

1988

1989

1990

1991

1992

1993

1994

1995

1996

1997

1998

1999

2000

2001

2002

2003

2004

2005

2006

2007

2008

2009

2010

2011

2012

2013

2014

2015

2016

2017

2018

2019

2020

2021

2022

2023

2024

2025

2026

2027

2028

2029

2030

2031

2032

2033

2034

2035

2036

2037

2038

2039

2040

2041

2042

2043

2044

2045

2046

2047

2048

2049

2050

2051

2052

2053

2054

2055

2056

2057

2058

2059

2060

2061

2062

2063

2064

2065

2066

2067

2068

2069

2070

2071

2072

2073

2074

2075

2076

2077

2078

2079

2080

2081

2082

2083

2084

2085

2086

2087

2088

2089

2090

2091

2092

2093

2094

2095

2096

2097

2098

2099

2100

2101

2102

2103

2104

2105

2106

2107

2108

2109

2110

2111

2112

2113

2114

2115

2116

2117

2118

2119

2120

2121

2122

2123

2124

2125

2126

2127

2128

2129

2130

2131

2132

2133

2134

2135

2136

2137

2138

2139

2140

2141

2142

2143

2144

2145

2146

2147

2148

2149

2150

2151

2152

2153

2154

2155

2156

2157

2158

2159

2160

2161

2162

2163

2164

2165

2166

2167

2168

2169

2170

2171

2172

2173

2174

2175

2176

2177

2178

2179

2180

2181

2182

2183

2184

2185

2186

2187

2188

2189

2190

2191

2192

2193

2194

2195

2196

2197

2198

2199

2200

2201

2202

2203

2204

2205

2206

2207

2208

2209

2210

2211

2212

2213

2214

2215

2216

2217

2218

2219

2220

2221

2

**Table 2.** Results of partial least squares regressions for fungal richness against environmental variables. Positive relationships are underlined; negative relationships are written in italics. \*\*\* indicates  $P < 0.001$ , \*\*  $0.001 > P < 0.01$ , \*  $0.01 > P < 0.05$ , blank indicates  $P > 0.05$ .

<b>Soil and environmental variables</b>	<b>Fungi (ITS)</b>	<b>Fungi (18S)</b>
<b>Total C<sup>L</sup></b>	0.44	<i>1.03 (<math>R^2 = 0.38^{***}</math>)</i>
<b>Total N<sup>L</sup></b>	0.93	0.56
<b>C:N ratio<sup>S</sup></b>	<i>1.64 (<math>R^2 = 0.28^{***}</math>)</i>	<i>1.71 (<math>R^2 = 0.41^{***}</math>)</i>
<b>Total P<sup>S</sup></b>	0.70	0.87
<b>Organic matter (% LOI)<sup>L</sup></b>	<i>1.13 (<math>R^2 = 0.29^{***}</math>)</i>	<i>1.17 (<math>R^2 = 0.38^{***}</math>)</i>
<b>pH (CaCl<sub>2</sub>)</b>	<u>1.52 (<math>R^2 = 0.23^{***}</math>)</u>	<u>1.55 (<math>R^2 = 0.37^{***}</math>)</u>
<b>Soil water repellency<sup>L</sup></b>	<i>1.23 (<math>R^2 = 0.13^{***}</math>)</i>	0.82
<b>Volumetric water content (m<sup>3</sup>/m<sup>3</sup>)</b>	0.60	0.70
<b>Rock volume (mL)</b>	0.64	0.43
<b>Bulk density (g/cm<sup>3</sup>)</b>	<u>1.41 (<math>R^2 = 0.29^{***}</math>)</u>	<u>1.33 (<math>R^2 = 0.41^{***}</math>)</u>
<b>Clay content (%)<sup>A</sup></b>	0.84	<u>1.19 (<math>R^2 = 0.11^{***}</math>)</u>
<b>Sand content (%)<sup>A</sup></b>	0.6	<i>1.11 (<math>R^2 = 0.1^{***}</math>)</i>
<b>Elevation (m)</b>	<i>1.68 (<math>R^2 = 0.22^{***}</math>)</i>	<i>1.83 (<math>R^2 = 0.41^{***}</math>)</i>
<b>Mean annual precipitation (mL)</b>	<i>1.44 (<math>R^2 = 0.18^{***}</math>)</i>	<i>1.52 (<math>R^2 = 0.27^{***}</math>)</i>
<b>Temperature (°C)</b>	0.56	0.52

Note: <sup>A</sup> denotes Aitchison's log<sub>10</sub>-ratio transformation; <sup>L</sup> denotes log<sub>10</sub>-transformation; <sup>S</sup> denotes square-root-transformation.

**Table 3.** Summary of relationships amongst environmental factors and fungal communities based on ITS data. +/- signify the direction of association between each variable and respective NMDS axes.

Variable	Correlation		
	R <sup>2</sup>	Axis1	Axis2
pH (CaCl <sub>2</sub> )	0.6***	-	+
C:N ratio <sup>S</sup>	0.47***	+	-
Elevation (m)	0.41***	+	-
Volumetric water content (m <sup>3</sup> /m <sup>3</sup> )	0.41***	+	-
Mean annual precipitation (mL)	0.39***	+	-
Bulk density (g/cm <sup>3</sup> )	0.38***	-	+
Organic matter (% LOI) <sup>L</sup>	0.37***	+	-
Total C <sup>L</sup>	0.31***	+	-
Clay content (%) <sup>A</sup>	0.28***	-	+
Soil water repellency <sup>L</sup>	0.24***	+	-
Total N (%) <sup>L</sup>	0.21***	+	-
Sand content (%) <sup>A</sup>	0.19***	+	+
Total P (mg/kg) <sup>S</sup>	0.11***	-	-
Rock volume (mL)	0.07***	-	+
Temperature (°C)	0.04***	-	+

Note: <sup>A</sup> denotes Aitchison's log<sub>10</sub>-ratio transformation; <sup>L</sup> denotes log<sub>10</sub>-transformation; <sup>S</sup> denotes square-root-transformation

**Table 4.** Summary of relationships amongst environmental factors and fungal communities based on 18S data. +/- signify the direction of association between each variable and respective NMDS axes.

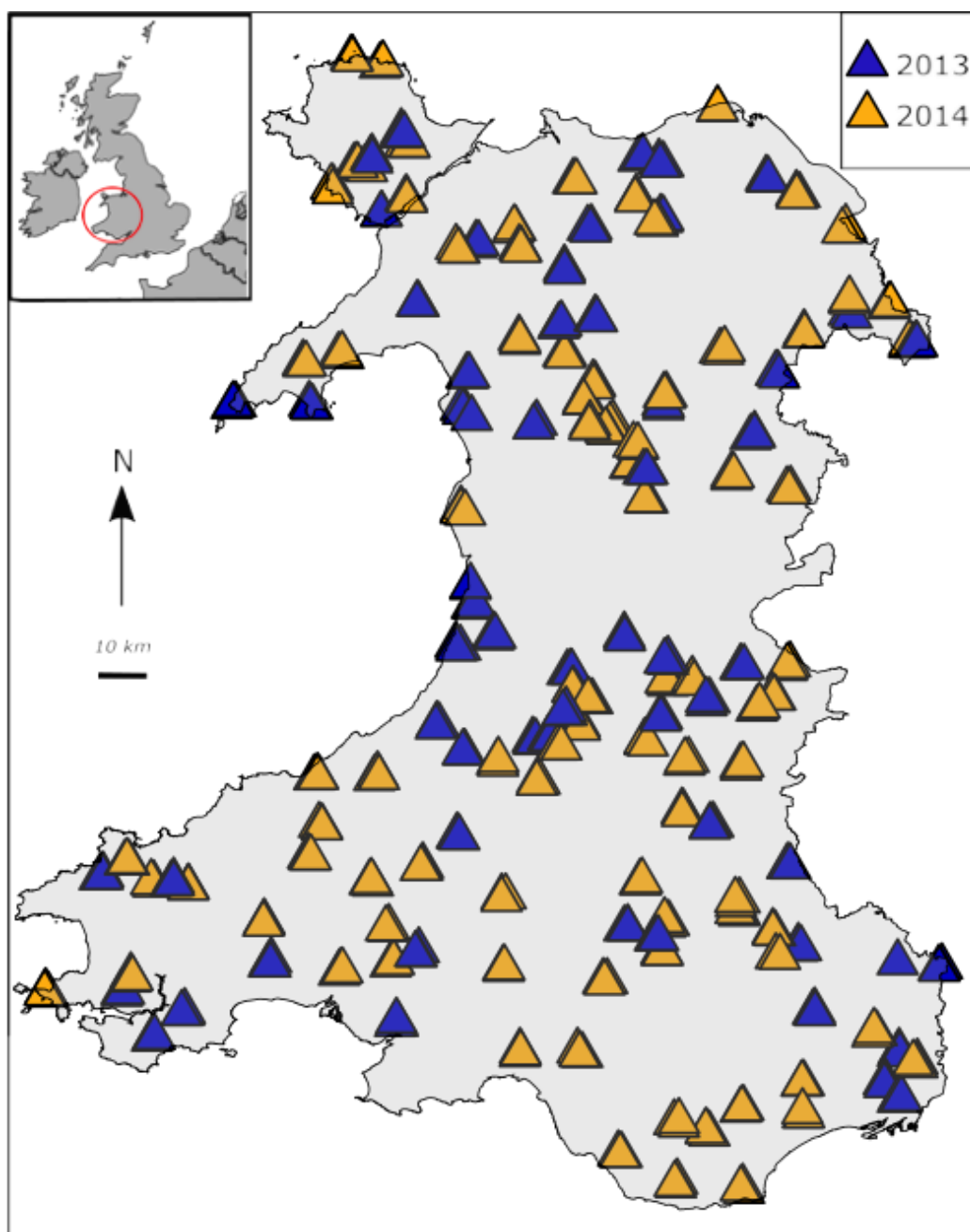
Variable	Correlation		
	R <sup>2</sup>	Axis1	Axis2
pH (CaCl <sub>2</sub> )	0.61 <sup>***</sup>	-	+
Elevation (m)	0.50 <sup>***</sup>	+	-
Mean annual precipitation (mL)	0.46 <sup>***</sup>	+	-
Volumetric water content (m <sup>3</sup> /m <sup>3</sup> )	0.45 <sup>***</sup>	+	-
C:N ratio <sup>S</sup>	0.43 <sup>***</sup>	+	+
Organic matter (% LOI) <sup>L</sup>	0.43 <sup>***</sup>	+	+
Bulk density (g/cm <sup>3</sup> )	0.39 <sup>***</sup>	-	-
Total C <sup>L</sup>	0.34 <sup>***</sup>	+	+
Clay content (%) <sup>A</sup>	0.30 <sup>***</sup>	-	+
Total N (%) <sup>L</sup>	0.28 <sup>***</sup>	+	-
Soil water repellency <sup>L</sup>	0.21 <sup>***</sup>	+	-
Sand content (%) <sup>A</sup>	0.14 <sup>***</sup>	+	+
Total P (mg/kg) <sup>S</sup>	0.10 <sup>***</sup>	-	-
Rock volume (mL)	0.06 <sup>***</sup>	-	+
Temperature (°C)	0.05 <sup>***</sup>	-	+

Note: <sup>A</sup> denotes Aitchison's log<sub>10</sub>-ratio transformation; <sup>L</sup> denotes log<sub>10</sub>-transformation; <sup>S</sup> denotes square-root-transformation

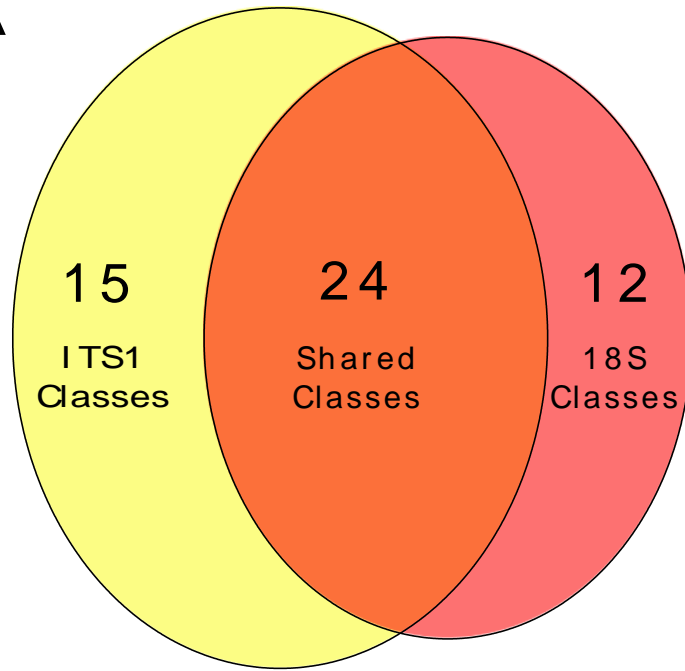
**Table 5.** Results of partial least squares regressions for richness of OTUs classified by trophic mode from FUNGUILD analyses against environmental variables. Positive relationships are underlined; negative relationships are written in italics. \*\*\* indicates  $P < 0.001$ , \*\*  $0.001 > P < 0.01$ , \*  $0.01 > P < 0.05$ , blank indicates  $P > 0.05$ .

Soil and environmental variables	Saprotrophs (ITS)	Saprotrophs (18S)	Pathotrophs (ITS)	Pathotrophs (18S)	Symbiotrophs (ITS)	Symbiotrophs (18S)
Total C (%) <sup>L</sup>	<i>1.1 (<math>R^2 = 0.24^{***}</math>)</i>	0.89	<i>1.07 (<math>R^2 = 0.17^{***}</math>)</i>	<i>1.0 (<math>R^2 = 0.25^{***}</math>)</i>	0.24	0.99
Total N (%) <sup>L</sup>	0.99	0.10	0.82	0.64	<i>1.17 (<math>R^2 = 0.02^{**}</math>)</i>	0.10
C:N ratio <sup>S</sup>	0.95	<i>2.31 (<math>R^2 = 0.28^{***}</math>)</i>	<i>1.22 (<math>R^2 = 0.16^{***}</math>)</i>	<i>1.41 (<math>R^2 = 0.25^{***}</math>)</i>	<u>1.69 (<math>R^2 = 0.01^*</math>)</u>	<i>2.47 (<math>R^2 = 0.34^{***}</math>)</i>
Total P (mg/kg) <sup>S</sup>	0.07	0.86	0.75	0.75	1.38	<u>1.31 (<math>R^2 = 0.02^*</math>)</u>
Organic matter (% LOI) <sup>L</sup>	<i>1.36 (<math>R^2 = 0.28^{***}</math>)</i>	<i>1.02 (<math>R^2 = 0.24^{***}</math>)</i>	<i>1.38 (<math>R^2 = 0.21^{***}</math>)</i>	<i>1.16 (<math>R^2 = 0.28^{***}</math>)</i>	0.37	0.92
pH (CaCl <sub>2</sub> )	<u>1.34 (<math>R^2 = 0.21^{***}</math>)</u>	<u>1.27 (<math>R^2 = 0.14^{***}</math>)</u>	<u>1.4 (<math>R^2 = 0.16^{***}</math>)</u>	<u>1.98 (<math>R^2 = 0.4^{***}</math>)</u>	<i>2.35 (<math>R^2 = 0.05^{***}</math>)</i>	<u>1.45 (<math>R^2 = 0.2^{***}</math>)</u>
Soil water repellency <sup>L</sup>	<i>1.28 (<math>R^2 = 0.15^{***}</math>)</i>	0.36	0.84	0.98	0.3	0.62
Volumetric water content (m <sup>3</sup> /m <sup>3</sup> )	<i>1.46 (<math>R^2 = 0.22^{***}</math>)</i>	0.56	<i>1.38 (<math>R^2 = 0.17^{***}</math>)</i>	0.99	<i>1.42 (<math>R^2 = 0.05^{***}</math>)</i>	0.40
Rock volume (mL)	0.68	0.06	0.8	0.59	<u>1.09 (<math>R^2 = 0.02^{**}</math>)</u>	0.10
Bulk density (g/cm <sup>3</sup> )	<u>1.42 (<math>R^2 = 0.28^{***}</math>)</u>	<u>1.23 (<math>R^2 = 0.2^{***}</math>)</u>	<u>1.71 (<math>R^2 = 0.12^{***}</math>)</u>	<u>1.29 (<math>R^2 = 0.27^{***}</math>)</u>	0.51	<u>1.48 (<math>R^2 = 0.26^{***}</math>)</u>
Clay content (%) <sup>A</sup>	0.71	0.74	0.90	<u>1.17 (<math>R^2 = 0.1^{***}</math>)</u>	0.49	<u>1.05 (<math>R^2 = 0.03^{**}</math>)</u>
Sand content (%) <sup>A</sup>	0.18	<i>1.71 (<math>R^2 = 0.05^{***}</math>)</i>	0.05	0.32	0.21	<u>1.63 (<math>R^2 = 0.08^{***}</math>)</u>
Elevation (m)	<i>1.58 (<math>R^2 = 0.25^{***}</math>)</i>	<i>1.13 (<math>R^2 = 0.13^{***}</math>)</i>	<i>1.6 (<math>R^2 = 0.19^{***}</math>)</i>	<i>1.98 (<math>R^2 = 0.39^{***}</math>)</i>	0.37	<u>1.07 (<math>R^2 = 0.17^{***}</math>)</u>
Mean annual precipitation (mL)	<i>1.45 (<math>R^2 = 0.23^{***}</math>)</i>	0.81	<i>1.38 (<math>R^2 = 0.16^{***}</math>)</i>	<i>1.49 (<math>R^2 = 0.24^{***}</math>)</i>	0.00	0.69
Temperature (°C)	0.09	0.49	0.21	0.43	<u>1.17 (<math>R^2 = 0.01^*</math>)</u>	0.53

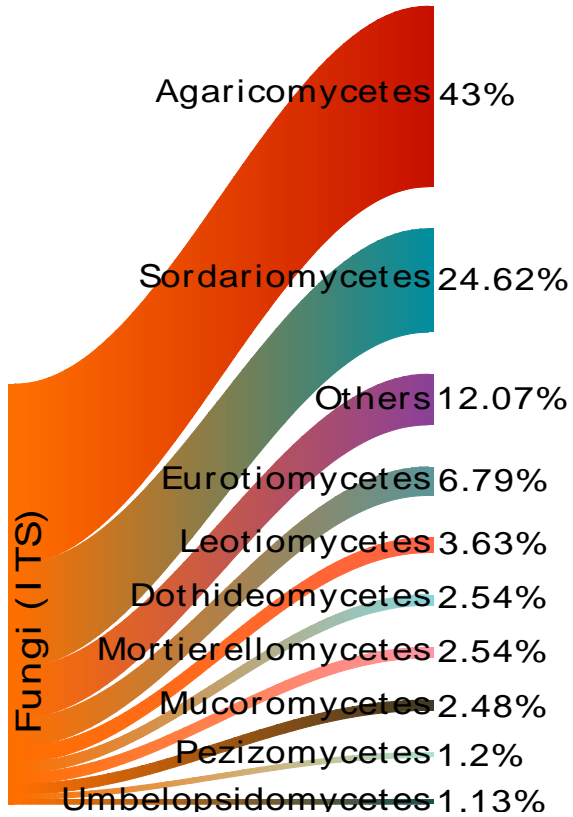
Note: <sup>A</sup> denotes Aitchison's log<sub>10</sub>-ratio transformation; <sup>L</sup> denotes log<sub>10</sub>-transformation; <sup>S</sup> denotes square-root-transformation



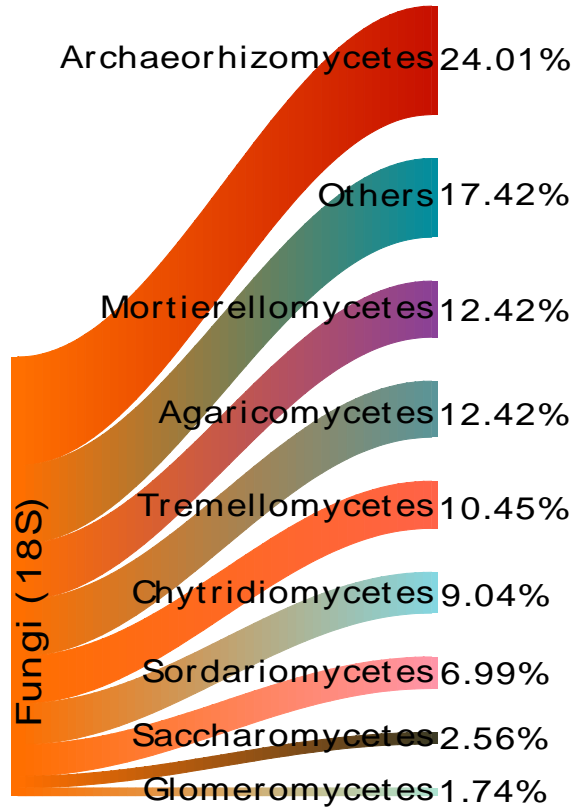
A

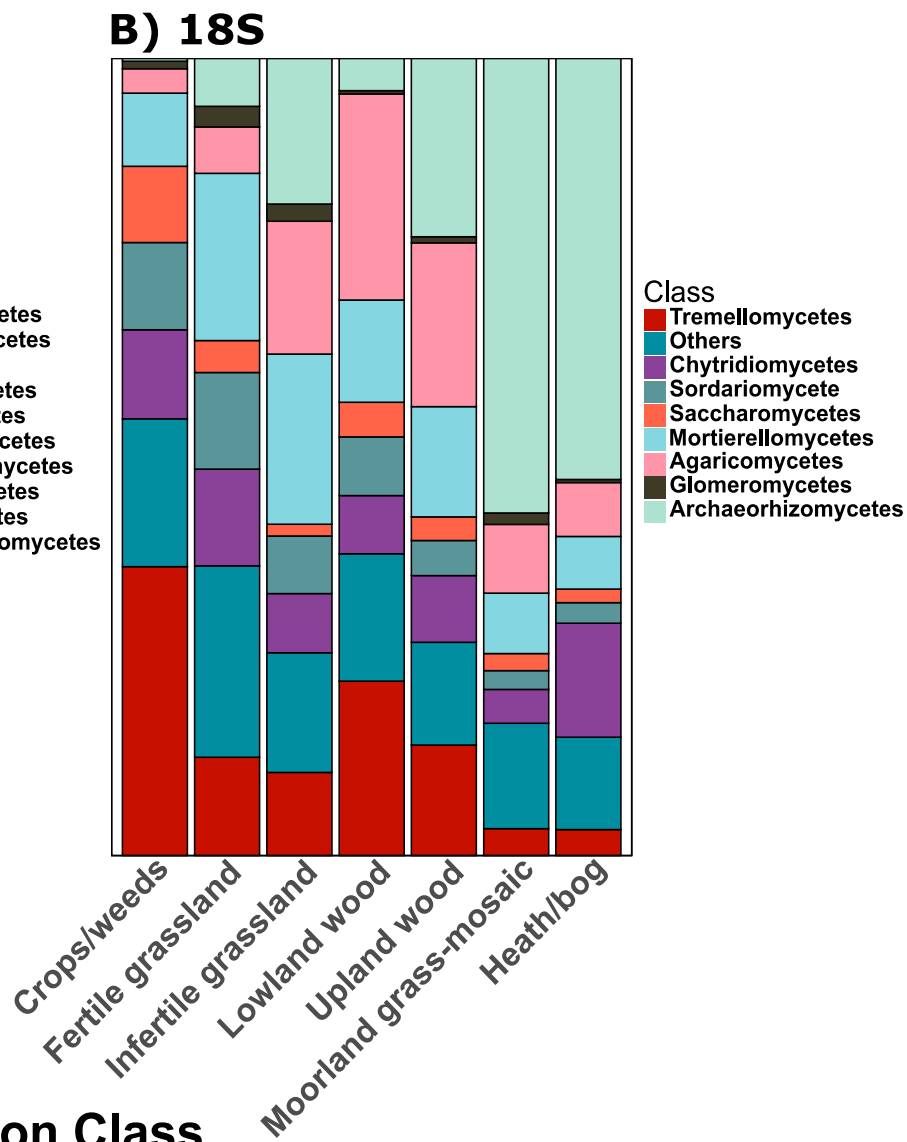
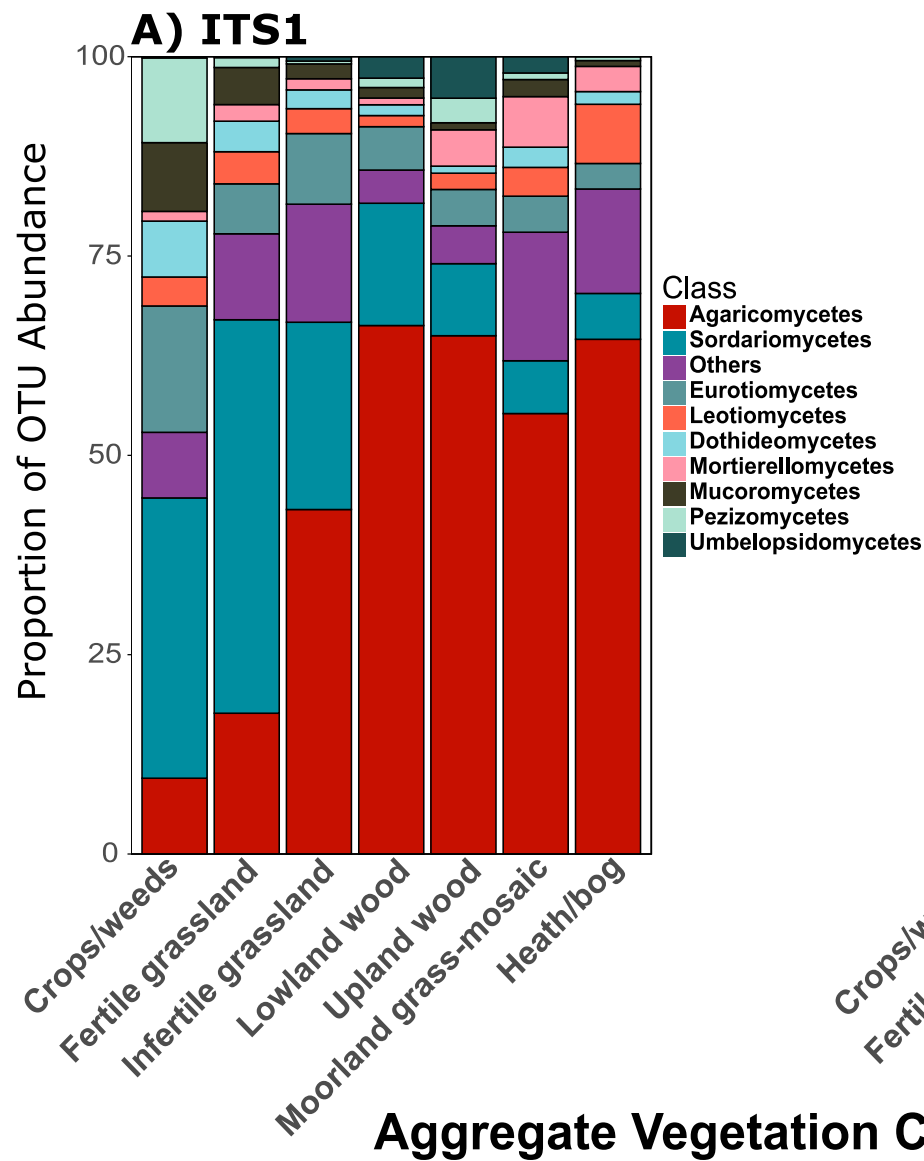


B

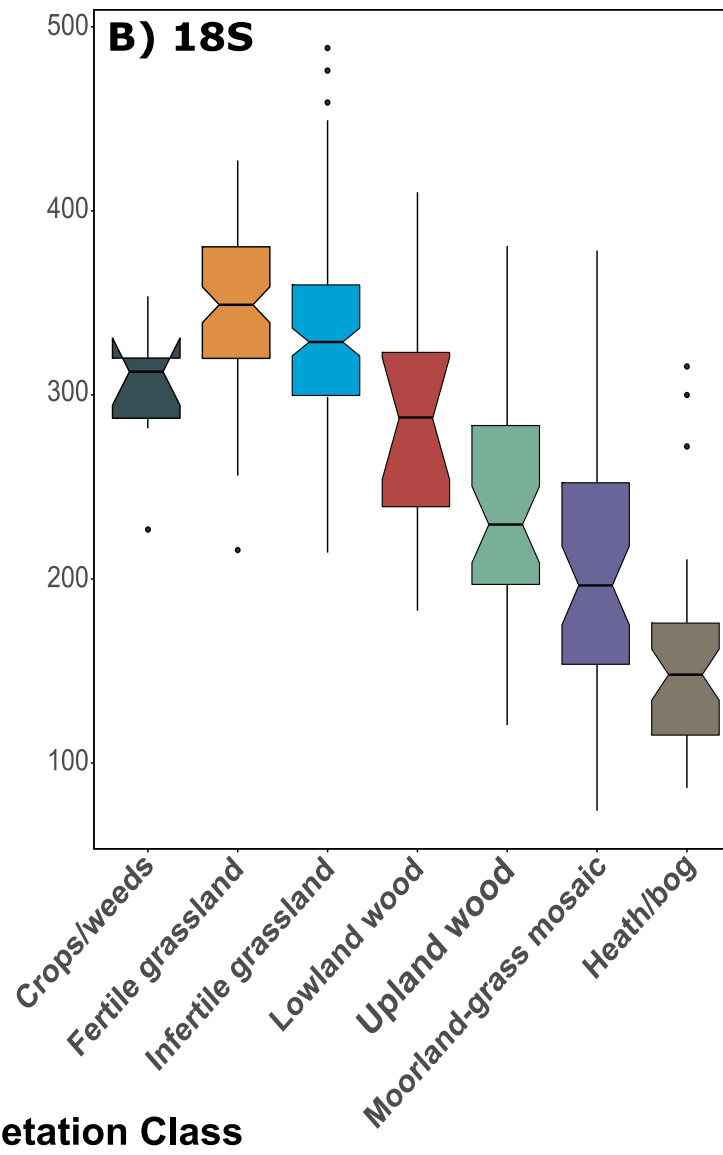
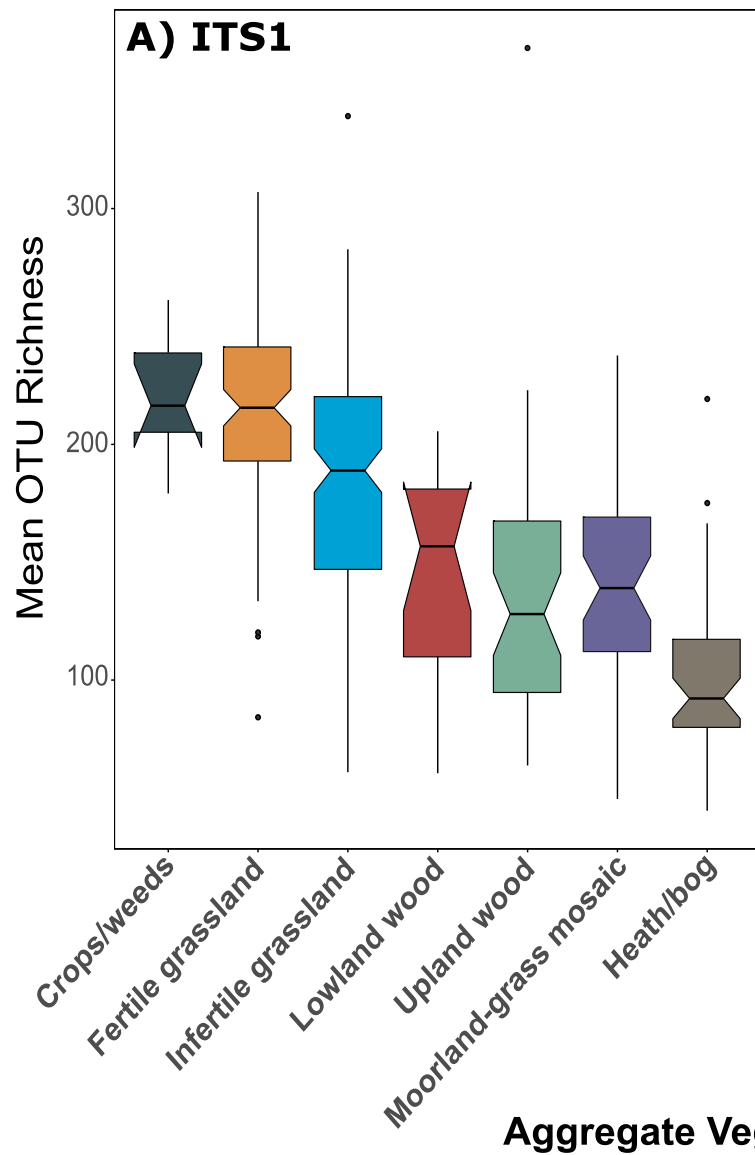


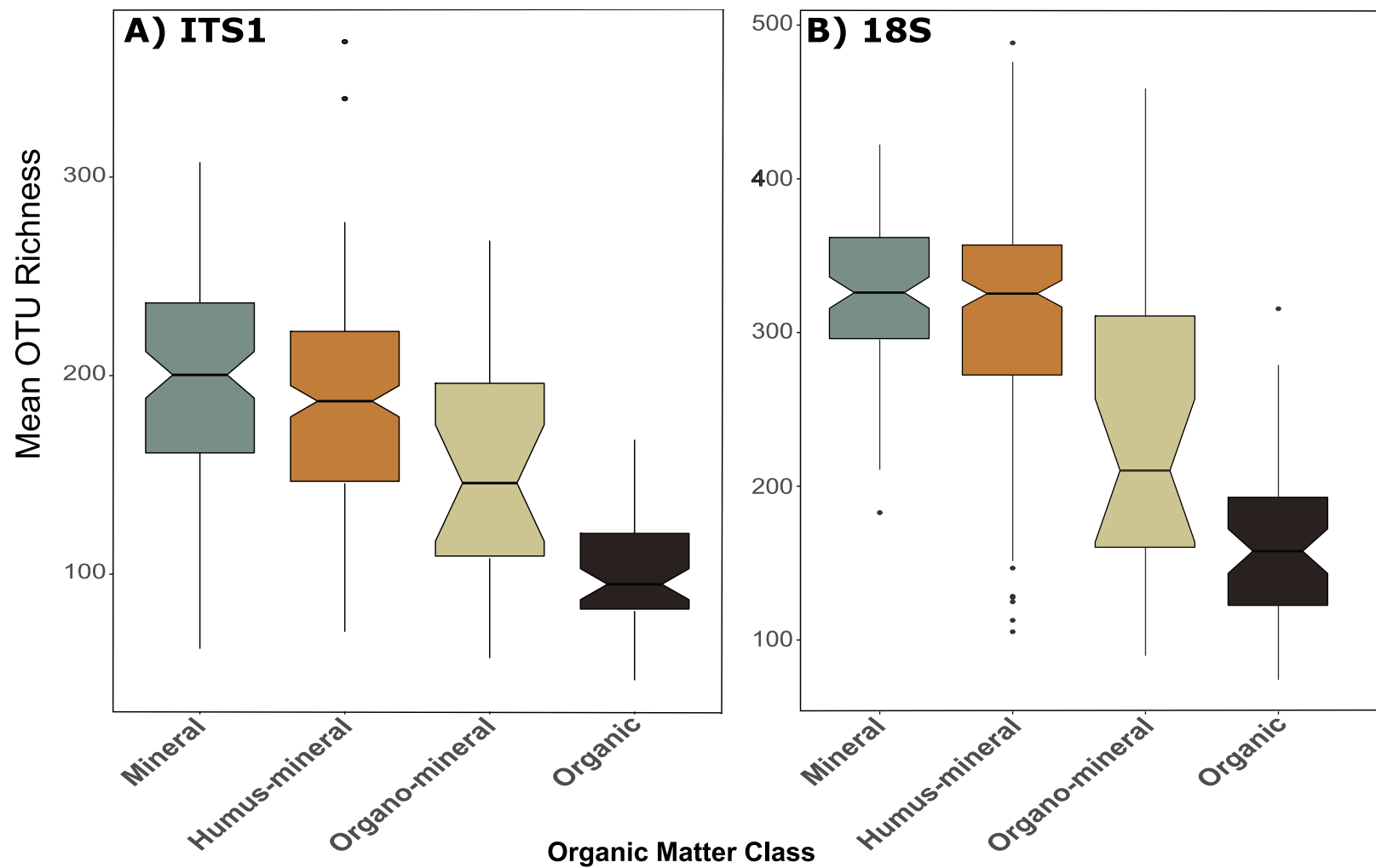
C



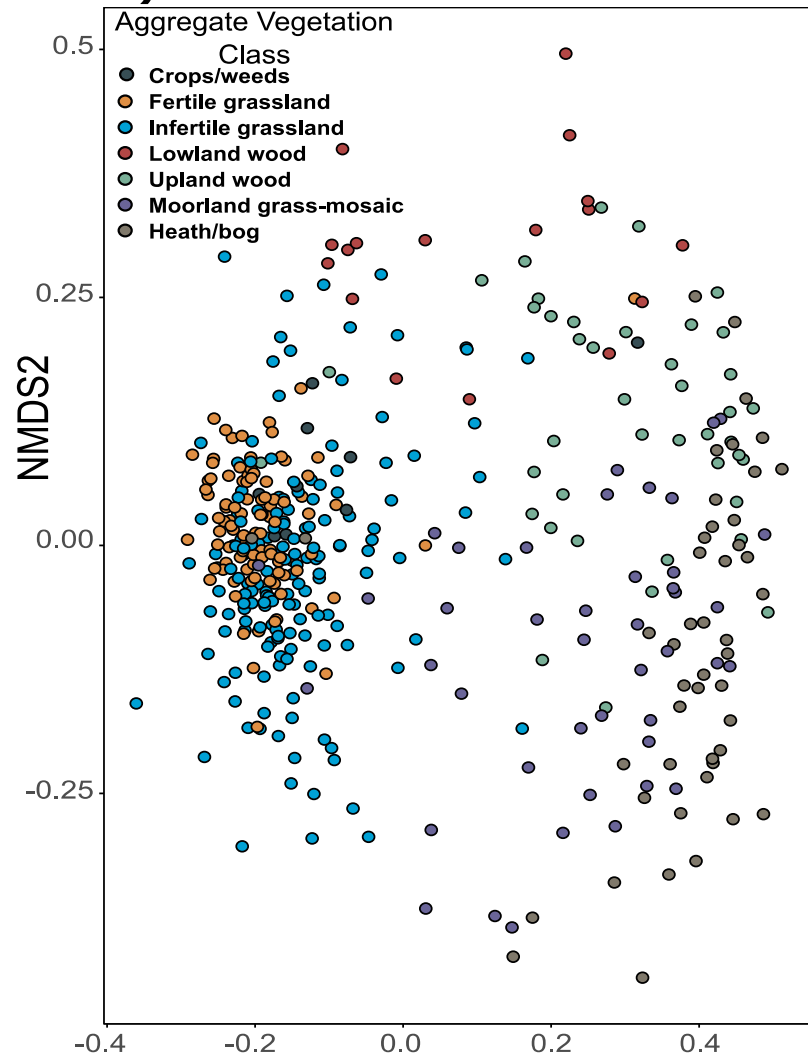




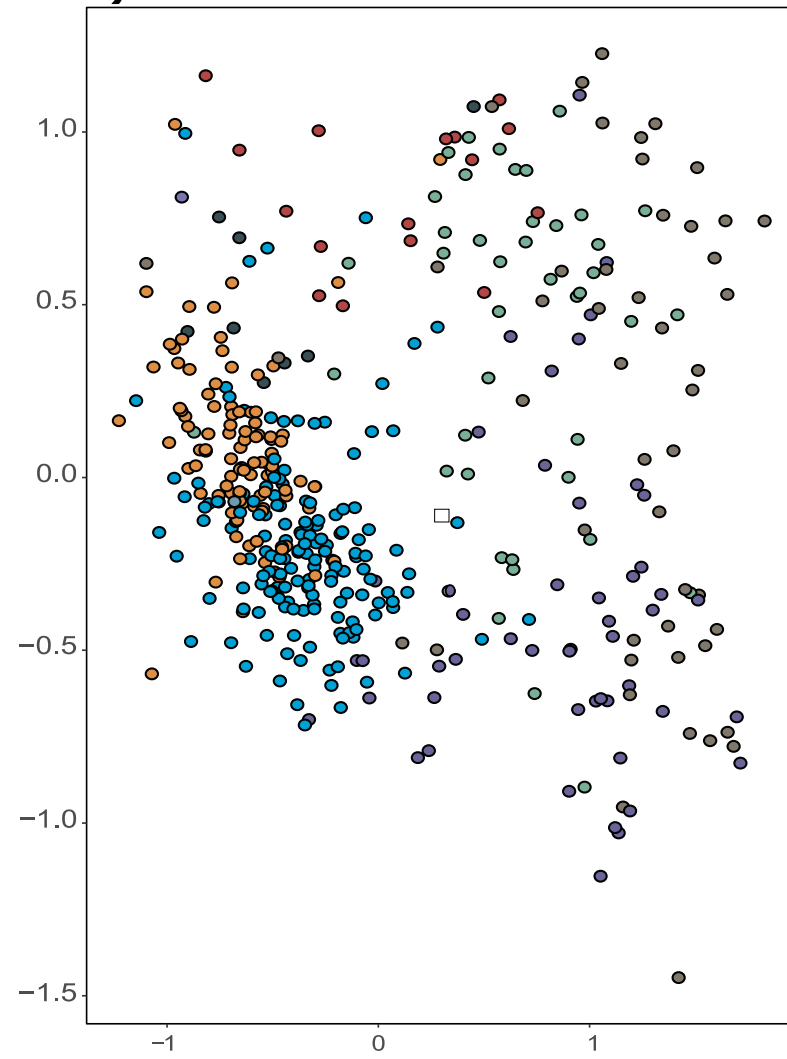


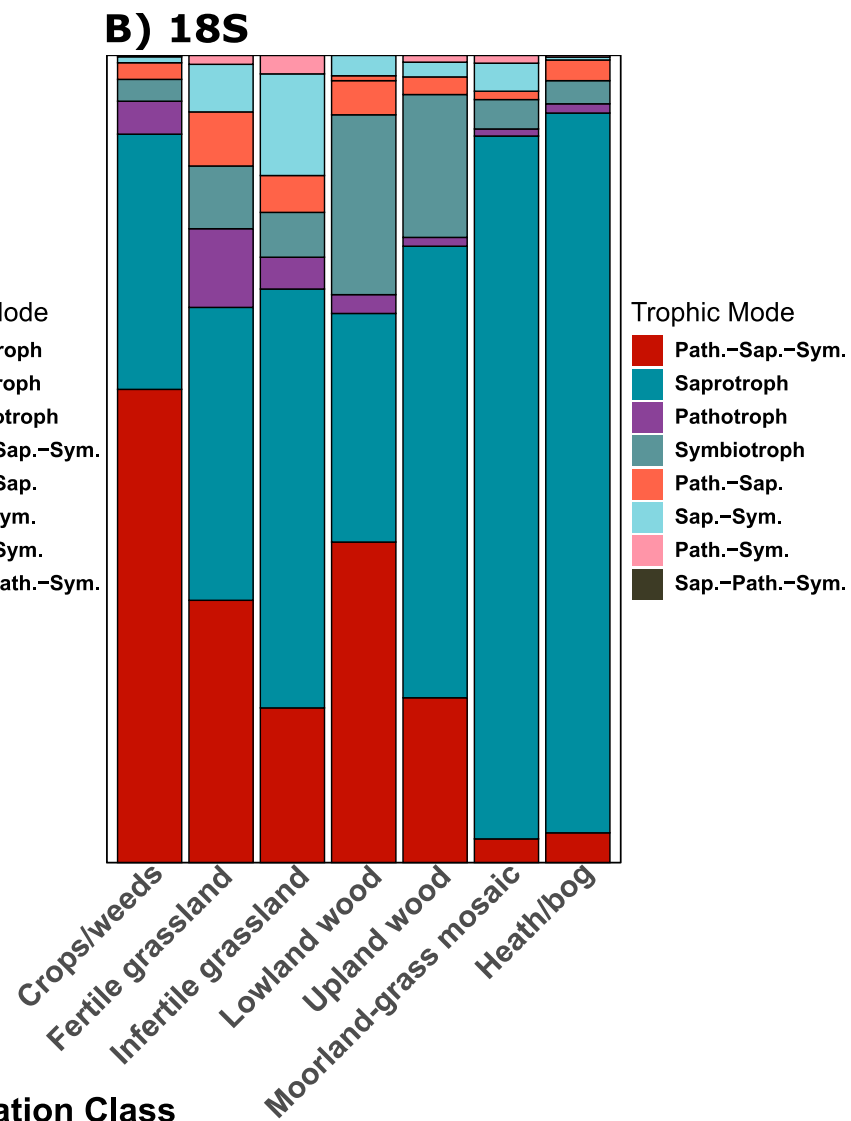
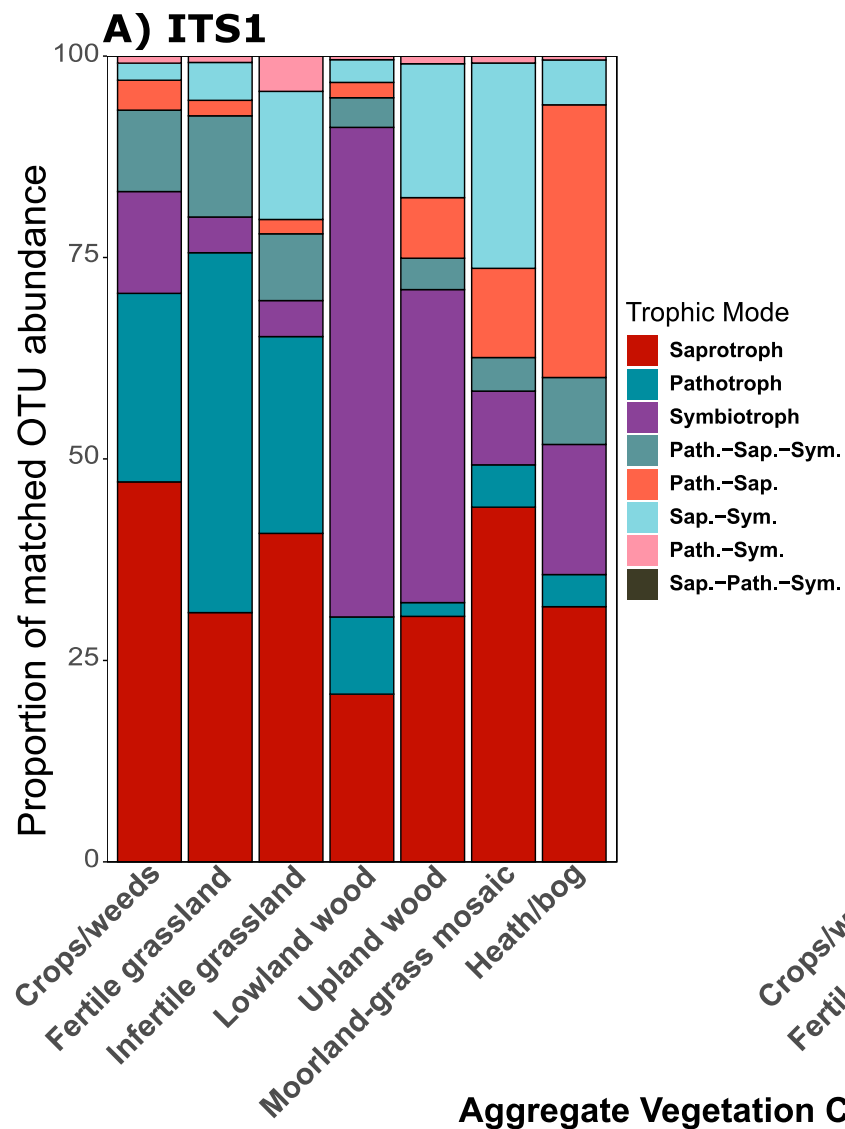


### A) ITS1

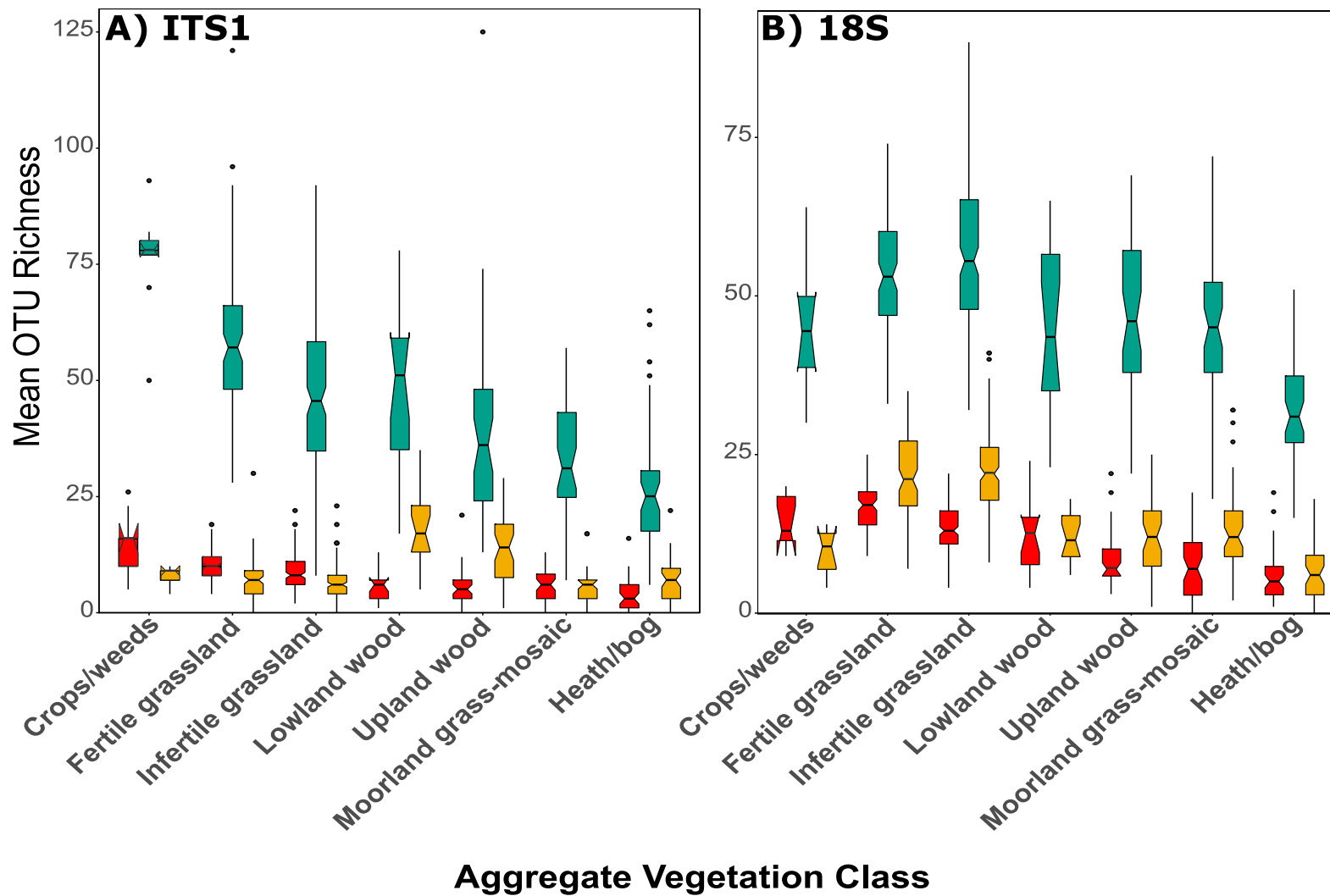


### B) 18S





Trophic Mode   Pathotroph   Saprotroph   Symbiotroph



Trophic Mode   Pathotroph   Saprotroph   Symbiotroph

